

UNIVERSITÉ DU QUÉBEC

MÉMOIRE SOUS FORME D'ARTICLES SCIENTIFIQUES
PRÉSENTÉ À
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN SCIENCES DE L'ENVIRONNEMENT

PAR
CHRISTINE GAUDREAU, B. Sc. (Biologie médicale)

IDENTIFICATION MOLÉCULAIRE DES SIMULIES
(DIPTERA : SIMULIIDAE) DU QUÉBEC

SEPTEMBRE 2005

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire ou de cette thèse a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire ou de sa thèse.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire ou cette thèse. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire ou de cette thèse requiert son autorisation.

AVANT-PROPOS

i) CONFORMITÉ DU MANUSCRIT

Conformément aux articles D45 à D47 du règlement des études de cycles supérieurs de l'Université du Québec à Trois-Rivières, il est possible de présenter les résultats obtenus dans le cadre du programme de 2^e cycle en sciences de l'Environnement sous forme d'article scientifique.

Le chapitre I de ce mémoire est composé d'un résumé substantiel du travail accompli, incluant une introduction englobant la problématique et les objectifs, les résultats, les conclusions et une liste de références.

Le chapitre II est constitué de la note scientifique décrivant la méthode d'extraction élaborée pour les fins de nos expérimentations; elle a pour titre «PCR suitability of field-extracted DNA from small invertebrates ». Les auteurs sont Christine Gaudreau, Bernard LaRue et Guy Charpentier. Le manuscrit sera soumis dans les prochaines semaines à *Nucleic Acids Research* pour fin de publication. Les instructions relatives à la publication d'articles dans ce journal se retrouvent à l'Annexe I.

Le chapitre III, qui englobe les résultats principaux accumulés durant ces deux années de travail et de ce fait constitue l'article scientifique principal, s'intitule « Characterization of black flies (Diptera: Simuliidae) from Quebec (Canada) by restriction-PCR ». Les auteurs sont les mêmes que ceux de la note scientifique, mais le présent article sera soumis à *Journal of Medical Entomology*. Les instructions relatives à la publication dans ce journal se retrouvent à l'Annexe II.

ii) REMERCIEMENTS

La réalisation d'un tel travail n'aurait pu avoir lieu sans la participation de certaines personnes que je tiens à mentionner.

Tout d'abord, je tiens à remercier mon directeur de recherche, le docteur Guy Charpentier, ainsi que mon co-directeur, le docteur Bernard LaRue, pour leurs judicieux conseils qui m'ont permis de mener à bien ce travail. Mille mercis!

Des remerciements sont également adressés aux techniciennes de laboratoire Micheline Boulianne, Diane Bégin et Nathalie Godbout pour leur générosité (matériels empruntés) et pour leurs précieux conseils. Je remercie également les docteurs Jacques Boisvert et Alain Maire pour les multiples services rendus.

Finalement, je voudrais remercier mes parents pour leurs encouragements constants et l'aide qu'ils ont pu m'apporter dans la réalisation d'un tel défi, sans oublier l'équipe du laboratoire, constituée de Lise Meilleur, Carole Charbonneau et Mylène St-Onge pour leur soutien scientifique ainsi que leur disponibilité ☺.

TABLE DES MATIÈRES

Avant propos	
i) Conformité du manuscrit.....	i
ii) Remerciements.....	ii
Table des matières.....	iii
Liste des tableaux.....	iv
Liste des figures.....	v
Symboles et abréviations.....	vii
Chapitre I Résumé français	
Introduction (problématique et objectif).....	2
Collection et identification morphologique des larves.....	4
Méthodes d'identification moléculaire.....	5
Méthodologie.....	5
Résultats.....	12
Conclusion.....	23
Bibliographie.....	25
Chapitre II Note scientifique	
Abstract.....	29
Note scientifique.....	30
References.....	32
Figures.....	33
Chapitre III Article scientifique	
Abstract.....	37
Article scientifique	
Introduction.....	39
Materials and methods	
Collection of black flies.....	43
DNA extraction.....	43
DNA amplification.....	43
Restriction fragment analysis.....	44
Heteroduplex assays (HDA).....	45
Results.....	46
Discussion.....	51
Acknowledgements.....	54
References.....	55
Table and figures.....	62
Annexe I	
Nucleic Acids Research submission guidelines.....	67
Annexe II	
Journal of Medical Entomology submission guidelines.....	70

LISTE DES TABLEAUX

Table 3.1	Black fly material used in the study.....	62
------------------	---	----

LISTES DES FIGURES

Figure 1.1 Critères d'identification morphologique basée sur la tête d'une larve de mouche noire.....	4
Figure 1.2 Espaceurs intergéniques transcrits de l'ADNr.....	8
Figure 1.3 Localisation relative et orientation transcriptionnelle des gènes mitochondriaux chez <i>D. yakuba</i>	9
Figure 1.4 Schéma explicatif de la technique des hétéroduplexes.....	12
Figure 1.5 Produits PCR du gènes mitochondrial COII.....	17
Figure 1.6 Produits PCR des gènes de l'ARN mitochondrial 12S.....	17
Figure 1.7 Produits PCR des gènes de l'ARN mitochondrial 16S.....	18
Figure 1.8 «Cluster» des gènes d'histones chez <i>D. melanogaster</i>	18
Figure 1.9 Sites de restriction spécifiques des endonucléases avec des cibles de 4 pb, à l'intérieur de l'espaceur H3-H4 de <i>D. melanogaster</i>	18
Figure 1.10 Séquences des amorces ciblant l'espaceur des gènes des histones H3 et H4.....	20
Figure 1.11 Taille de l'espaceur H3-H4.....	21
Figure 1.12 Schème de classification de 12 espèces de mouches noires pour les marqueurs ITS-1 et H3-H4 seuls ou combinés.....	22
Figure 1.13 Principe du PCR multiplexe.....	22
Figure 1.14 Principe du PCR sélectif de l'espèce.....	24
Figure 2.1 Agarose gel electrophoresis of black fly larva DNA extracted through various methods.....	33

Figure 2.2 Effect of dilution of black fly DNA on the PCR signal of ITS-1.....	34
Figure 2.3 Effect of storage temperature of field-extracted DNA on the PCR signal.....	34
Figure 3.1 ITS-1 profile on agarose and polyacrylamide gels.....	63
Figure 3.2 Dde I, Alu I and Rsa I fingerprints.....	64
Figure 3.3 HDA profiles from Dde I digest of ITS-1.....	65
Figure 3.4 Pathway for the classification of 12 black fly species.....	66

SYMBOLES ET ABRÉVIATIONS

Cette section comporte une description bilingue des abréviations utilisées dans les portions francophones et anglophones du manuscrit. Afin de simplifier la consultation de cette liste, l'ordre alphabétique des abréviations fut préféré à l'ordre d'apparition dans le texte.

~ : environ

λ : bactériophage lambda

* : multiplié

A: adénine

ADN (DNA): acide désoxyribonucléique (deoxyribonucleic acid)

ADNr (rDNA): ADN ribosomal (ribosomal DNA)

Ala: alanine

Arg: arginine

ARN (RNA): acide ribonucléique (ribonucleic acid)

ARNr (rRNA): ARN ribosomal (ribosomal RNA)

Asn: asparagine

C: cytosine

cm: centimètre

COII: sous-unité II de la cytochrome oxydase

dNTP: désoxyribonucléotides 5' triphosphatés

EDTA: acide éthylènediaminetétraacétique

EtOH: éthanol

G: guanine

Gln: glutamine

HDA: analyse des hétéroduplexes (heteroduplex assay)

I: désoxyinosine

ITS: espaceur interne transcrit (internal transcribed spacer)

ITS-1: espaceur transcrit 1 de l'ADNr nucléaire

ITS-2: espaceur transcrit 2 de l'ADNr nucléaire

Lys : lysine

kb: kilobase

mg: milligramme

min: minute

mL: millilitre

mM: millimolaire

mtDNA: ADN mitochondrial (mitochondrial DNA)

M: concentration molaire

Met: méthionine

ng: nanogramme

pb (bp): paire de bases (base pair)

pmol: picomole
PCR: réaction en chaîne de la polymérase (polymerase chain reaction)
Phe: phénylalanine
Pro: proline
ProK: protéinase K
RAPD: «random amplified polymorphic DNA»
s: seconde
SDS: dodécyl sulfate de sodium
SSCP: polymorphisme conformationnel simple brin
T: thymine
Taq: ADN polymérase de *Thermus aquaticus*
Thr: thréonine
T_m: température de fusion (melting temperature)
U: unité enzymatique
UV: ultraviolet
μL: microlitre
μM: micromolaire
V: volts
VNTR: nombre variable de répétitions en tandem (variable number of tandem repeats)

CHAPITRE I

Résumé français

Introduction (problématique et objectif)

Les mouches noires (simulies) causent dans le monde entier des pertes économiques non négligeables au niveau de l'élevage du bétail et de la volaille et agissent dans les pays tropicaux comme vecteurs de maladies telles l'onchocercose. Au Canada, elles ont aussi, par leur effet de nuisance, un impact non négligeable sur le tourisme. Cependant, aucune transmission à l'homme d'organismes pathogènes ou de parasites n'a été répertoriée au Québec, excepté quelques réactions allergiques, parfois sévères, dues aux morsures.

Bien que plusieurs travaux scientifiques aient été effectués durant le dernier siècle, la taxinomie des simulies, principalement en zone néarctique, repose encore largement sur la morphologie des larves et des adultes; il n'existe à peu près pas de clé d'identification pour les œufs ou les nymphes. Cependant, cette famille comprend relativement peu d'espèces comparativement aux autres familles de diptères et des types de gîtes similaires (milieux lotiques) associés à un mode de développement spécialisé des stades immatures lui confèrent une grande uniformité morphologique, réduisant ainsi le nombre de caractères distinctifs permettant de séparer les espèces. L'identification à l'espèce de ses représentants demeure ardue et ne donne pas toujours des réponses précises. De plus, nous avons souvent affaire à des complexes d'espèces-sœurs pratiquement identiques au plan morphologique, ce qui nous laisse facilement croire que les études morphologiques, bien que couramment utilisées, ne sont peut-être pas aussi exactes qu'on le voudrait. Plus récemment, les approches cytogénétiques et moléculaires se sont ajoutées aux clés morphologiques, mais aucune méthode d'identification existante ne permet encore une identification simple, rapide et s'appliquant à un large éventail d'espèces.

Les méthodes cytogénétiques, basées principalement sur l'examen des chromosomes géants des glandes salivaires chez la larve, s'appliquent surtout aux derniers stades larvaires (Xiong et Kocher, 1991). Elles ne permettent l'identification des femelles adultes, importantes épidémiologiquement (Krüger *et al.*, 2000), que lorsqu'il y a présence des tubes de Malpighi et des cellules nutritives des ovaires. Bien que la

cytogénétique ait montré que plusieurs espèces considérées jusqu'à récemment comme des espèces uniques étaient en fait des complexes d'espèces jumelles mais distinctes, elle demeure techniquement difficile à utiliser et requiert énormément de temps d'essais et de larves avant d'obtenir des résultats satisfaisants.

Complétant la morphologie et la cytogénétique, le typage de l'ADN peut s'avérer très utile. Depuis 1991, de plus en plus d'études basées sur cette approche ont été publiées, mettant même de côté les deux autres méthodes. Le typage d'ADN est principalement basé sur l'analyse des produits d'amplification de la réaction en chaîne de la polymérase (PCR). Ceci inclut le séquençage de l'ADN amplifié ou, à un niveau plus global, la détermination de la taille du produit PCR, l'identification de polymorphismes des sites de restriction et l'analyse des hétéroduplexes (HDA) ou du polymorphisme conformationnel simple brin (SSCP).

Bien que dans de rares cas les méthodes moléculaires ne puissent être appliquées pour résoudre un cas spécifique, elles sont, contrairement aux autres méthodes, valables pour les deux sexes et pour toutes les étapes du cycle de vie des insectes (Tang *et al.*, 1995a, 1995b et 1998). Finalement, nous croyons que les marqueurs moléculaires fourniront une image plus exacte des distances génétiques impliquées et établiront de façon claire la distinction entre les espèces difficiles à différencier selon les critères usuels. C'est donc sur ce choix que reposait le point de départ pour l'établissement d'une clé d'identification moléculaire des similies situées en zone néarctique.

En fonction d'une approche expérimentale en deux parties, soient d'abord la collection et l'identification morphologique des larves, puis ensuite la mise au point d'une clé moléculaire, nous avons formulé les quatre objectifs suivants :

1. Constituer une banque de spécimens d'un nombre maximum d'espèces de similies provenant de divers sites et ce, tout au long de l'année.

2. Identifier grâce aux clés morphologiques un nombre appréciable de similies, principalement des larves recueillies à travers le Québec.
3. Établir une méthode d'extraction de l'ADN réalisable sur le terrain et nous permettant de préserver celui-ci intact pour quelques jours à température ambiante.
4. Soumettre l'ADN extrait à différentes techniques moléculaires dans le but d'en faire une clé universelle d'identification moléculaire.

Collection et identification morphologique des larves.

Dans un premier temps, l'établissement d'une telle clé moléculaire nécessite un large éventail d'espèces distinctes. Différentes décharges de lacs ont été échantillonnées tout au long de l'année un peu partout au Québec (L'Assomption, Granby, St-Hippolyte) et à Ottawa (Ontario) avec la majorité des espèces provenant de la région de la Mauricie. Ainsi, en variant nos aires et nos dates d'échantillonnages, nous avons pu porter finalement à 11 le nombre d'espèces distinctes (objectif 1). Les larves récoltées étaient directement fixées dans de l'éthanol 95% et conservées à 4°C pour ensuite être identifiées à l'espèce sous binoculaire grâce aux clés d'identification visuelle (fig. 1.1) de Wood *et al.* (1963) ainsi qu'à celles de Adler *et al.* (2004). Cette diversité au niveau des clés utilisées avait pour but d'accroître le nombre de critères permettant dans certains cas une meilleure identification.

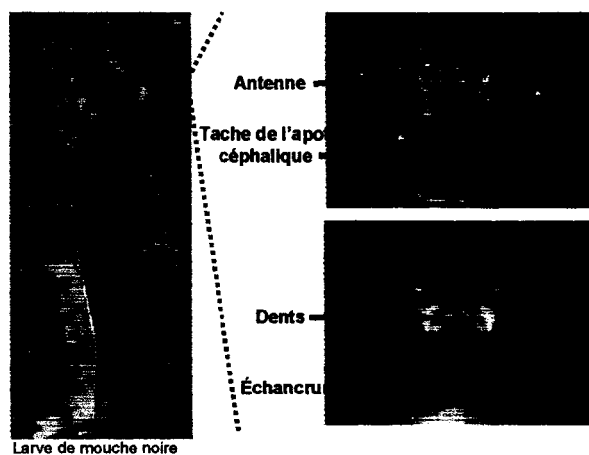


Figure 1.1 Critères d'identification morphologique basée sur la tête d'une larve de mouche noire. Figures tirées de Craig (2004).

Méthodes d'identification moléculaire.

Pour l'établissement de la clé moléculaire, différentes techniques ont été mises à profit pour typer onze espèces morphologiques préalablement identifiées. Comme ces méthodes devaient être rapides et mettre en évidence des différences interspécifiques, mais non des différences intraspécifiques, on a exclu au départ le typage par RAPDs ou minisatellites. De plus, bien qu'on doive écarter le séquençage comme méthode de routine, des données de séquence pourraient être éventuellement requises s'il devenait nécessaire de créer des amorces spécifiques en vue d'exécuter ultérieurement un PCR sélectif de l'espèce. Pour la suite, la stratégie était la suivante : repérer de courtes (<2 kb) régions variables de séquence et de longueur, mais encadrées par des séquences conservées constituant des cibles pour des amorces PCR d'application universelle. Nous avons donc utilisé cette variabilité comme critère d'identification de l'espèce et nous avons cherché à combiner les informations de plusieurs cibles de ce type afin d'aboutir à une clé systématique comparable à celle issue des critères morphologiques.

MÉTHODOLOGIE

La fixation de spécimens frais dans l'éthanol 95% et leur conservation à 4°C constituent le meilleur choix pour la poursuite de nos expériences; ces conditions semblent fournir une meilleure qualité de l'ADN qu'à de plus basses concentrations d'éthanol ou à de plus hautes températures (Post *et al.*, 1993). De plus, ce traitement diminue la distorsion et le changement de couleur des larves et entraîne généralement l'extension des plumeaux céphaliques, ce qui s'avère très utile lors de l'identification morphologique (Wood *et al.*, 1963).

La littérature répertorie un très grand nombre de méthodes d'extraction de l'ADN. Même si plusieurs techniques ont été comparées entre elles (Steiner *et al.*, 1995) quant au travail requis, à la complexité et au coût, il n'en demeure pas moins que l'on ne retrouve aucune uniformité au niveau de la littérature. Par ailleurs, les trousseaux commerciaux basés sur des résines de silice poreuse sont en train de supplanter

beaucoup de méthodes manuelles plus pénibles. Cependant, ces trousseaux capables de fournir de façon routinière de l'ADN de «qualité PCR» peuvent difficilement s'utiliser sur le terrain, où on préférera plutôt une méthode de faible technologie, mais robuste, pour extraire l'ADN et le conserver sans précautions spéciales jusqu'au retour en laboratoire où la procédure de purification sera complétée. Sur cette base, nous avons donc préalablement établi notre propre méthode d'extraction à partir de différents essais puis nous l'avons ensuite comparée à celle, dite «one step extraction», décrite dans l'article de Steiner *et al.* (1995) et présentée par ces auteurs comme une des plus simples existantes. Évidemment, pour les besoins de nos expérimentations, notre technique devait être en mesure d'extraire, dans des conditions de terrain, l'ADN de petits spécimens (ex : drosophile, simule ou parasites de simules tel des nématodes) tout en nous permettant de conserver les extraits dans le tampon d'extraction pour quelques jours à température ambiante.

Chacune des techniques d'extraction citées dans la littérature comporte son lot d'avantages et d'inconvénients. Quoi qu'il en soit, cette étape devra toutefois répondre aux critères suivants: être applicable sur un large éventail d'espèces, être simple à réaliser et peu dispendieuse tout en diminuant le nombre de manipulations et de matériels requis, le tout sans générer des matières toxiques dommageables à l'environnement (Steiner *et al.*, 1995). Puis, comme les expérimentateurs n'ont pas toujours accès à des conditions favorables pour conserver sur le terrain même les spécimens recueillis, un bon tampon d'extraction permettant de préserver l'ADN intact quelques jours à température ambiante serait approprié. Quant à sa composition, ce tampon d'extraction doit contenir à titre d'agent préservatif soit un détergent comme le SDS, soit de l'EDTA. De plus, un autre critère rarement mentionné dans la littérature semble primordial quant au succès du PCR, à savoir l'élimination d'impuretés présentes dans les extraits bruts d'ADN et qui agissent comme inhibiteurs de l'ADN polymérase. Ce problème d'inhibiteur pourrait être contourné en diluant l'ADN extrait tout en espérant conserver la capacité d'amplification, mais, vu la faible quantité d'ADN présente dans nos petits spécimens, on risque alors de perdre le signal PCR. Une fois extrait l'ADN brut, une purification supplémentaire semble désirable. Deux méthodes de

purification peuvent être appliquées dans le cas présent : une précipitation à l'éthanol ou une purification sur résine de silice. Bien que la purification sur résine demande un peu plus de manipulations et soit plus coûteuse que la précipitation à l'éthanol, la première élimine toutefois pratiquement tous les contaminants et est plus efficace que la précipitation. Une fois complétées l'extraction et la purification de l'ADN, il ne nous reste donc qu'à effectuer le choix des amorces servant à cibler une séquence spécifique sur l'ADN génomique et à déterminer les conditions requises pour le PCR, en particulier la quantité d'ADN qui, injectée dans une réaction d'amplification, donne le rendement optimal en produit PCR. Lors d'une amplification exécutée dans un volume de 25 μ L, cette quantité peut varier entre 12 et 0.06 ng tout en permettant de conserver un profil PCR reproductible et bien défini.

Comme mentionné plus tôt, la stratégie était de repérer de courtes régions variables de séquence et de longueur, encadrées par des séquences conservées constituant des cibles pour des amorces PCR. De plus, la taille de la région amplifiée ne doit pas excéder 2 kb, limite pratique du PCR conventionnel exécuté avec l'ADN polymérase de *Thermus aquaticus*. Pour ce faire, cinq marqueurs ont été choisis. Les régions que nous ciblerons seront les suivantes et les raisons justifiant leur choix seront décrites par la suite:

- 1) L'espaceur interne transcrit 1 de l'ADNr nucléaire (ITS-1) ; Tang *et al.*, 1998.
- 2) Gène d'ARNr mitochondrial 16S ; Tang *et al.*, 1998.
- 3) Gène d'ARNr mitochondrial 12S ; Tang *et al.*, 1998.
- 4) Gène mitochondrial de la sous-unité II de la cytochrome oxydase (COII) ; Tang *et al.*, 1998.
- 5) L'espaceur intergénique situé entre les gènes nucléaires des histones H3 et H4.

Le premier marqueur analysé fut l'espaceur transcrit 1 (ITS-1) localisé entre les gènes d'ARNr 5.8S et 18S. Deux raisons nous ont poussé à débiter avec celui-ci: tout d'abord, ces amorces conçues au départ par Tang *et al.* (1998) pour amplifier l'ADN de similies du complexe *Psilopelmia* fonctionnent tout aussi bien, vu la conservation de séquences, avec les lépidoptères et coléoptères, ce qui indique qu'on peut les appliquer à tout

génomique d'insecte. Ensuite, l'ITS-1 est une séquence destinée à l'élimination qui semble beaucoup plus variable que celle des gènes qui l'encadrent. L'ITS-1 est d'ailleurs reconnu comme une région avec un haut niveau de divergence évolutive et montrant parfois une variabilité dans sa structure et sa séquence, même entre espèces étroitement apparentées (Morales-Hojas *et al.*, 2002). De plus, l'opéron de l'ADNr est présent chez les diptères en plusieurs centaines de copies, presque toutes regroupées sous forme d'unités en tandem cartographiées sur un très petit nombre de sites chromosomiques (Hwang et Kim, 1999), ce qui facilite la détection PCR. D'ailleurs, grâce à ce marqueur amplifié, puis analysé à partir de ses profils de restriction, Iranpour *et al.* (2004) ont réussi à différencier avec succès 34 espèces de Tabanidae du Manitoba.

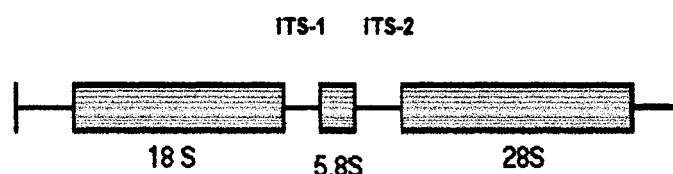


Figure 1.2 Espaces intergéniques transcrits de l'ADNr.

Les amorces utilisées pour caractériser le premier espaceur intergénique (ITS-1) ont été précédemment décrites par Tang *et al.* (1998) :

amorce 1 (extrémité droite du 18S) : 5' TGGTGCATGGCGTTCTTAC 3'

amorce 2 (extrémité gauche du 5.8S) : 5' GTCGATGTTTCATGTGTCCTGC 3'

D'autres marqueurs fréquemment utilisés, tels les gènes mitochondriaux COII, 12S et 16S pourraient aussi en principe servir à différencier les espèces non résolues par l'ITS-1. Cependant, selon nos résultats préliminaires, ces marqueurs semblent présenter moins de potentiel que ITS-1. À la suite d'une amplification réussie utilisant des amorces de séquences déjà publiées (Tang *et al.*, 1998), nous avons en effet obtenu des amplicons de tailles petites et uniformes offrant peu de pouvoir discriminant entre espèces. Les amorces généralement employées par les différents auteurs ciblent le gène et non un espaceur, réduisant ainsi l'aspect variabilité. De plus, deux autres problèmes potentiels peuvent être soulevés dans le cas des gènes mitochondriaux en général : le risque de

confusion, vu leur transmission maternelle exclusive, entre polymorphisme intraspécifique et variation interspécifique, ainsi que la création fréquente de pseudogènes nucléaires issus de la migration vers le noyau de gènes mitochondriaux. Dans ce dernier cas, cette migration conduirait à des conclusions erronées quant à l'identification à l'espèce (Pineau *et al.*, 2005). Malgré tout, ces marqueurs sont souvent employés lors des analyses phylogénétiques car ils sont hautement conservés entre phyla, sont présents en copies multiples et évoluent plus vite que l'ADN nucléaire (Morlais et Severson, 2002; Szalanski et Owens, 2003).

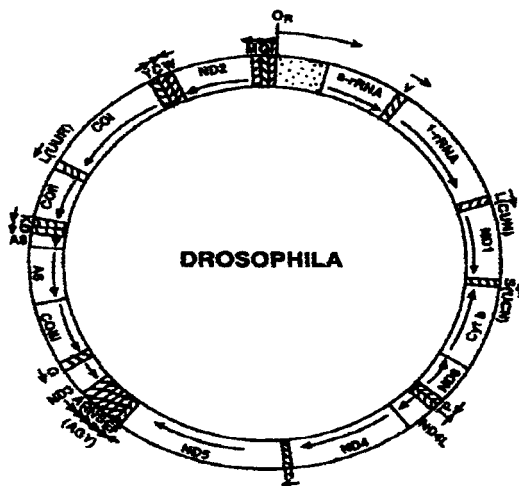


Figure 1.3 Localisation relative et orientation transcriptionnelle des gènes mitochondriaux chez *Drosophila yakuba* (Simon *et al.*, 1994).

Bien que Krüger *et al.* (2000) et Tang *et al.* (1995a et 1995b) aient trouvé dans des cas particuliers que le gène mitochondrial 16S était utile lors de l'identification à l'espèce (Mank *et al.*, 2004) et réussissait à différencier quelques espèces de mouches noires nord-américaines (Tang *et al.*, 1996), il reste que, dans nos expérimentations, aucun des marqueurs mitochondriaux utilisés n'a permis d'atteindre cet objectif.

Les gènes d'histones furent également choisis comme cibles potentielles car, tout comme ceux des espaceurs intergéniques des ARNr, ils font partie d'une famille de séquences répétées en tandem facilitant ainsi l'obtention d'un produit PCR dû à la multiplicité des cibles (Mank *et al.*, 2004; Kremer et Hennig, 1991). Chez le nématode *Caenorhabditis elegans*, le bloc regroupant les gènes de H1, H2A, H2B, H3 et H4 est dispersé à travers le génome en groupes d'une ou plusieurs copies (Rooney *et al.*, 2002).

tandis que chez *D. melanogaster*, ces gènes sont organisés en un seul type de répétition en tandem comprenant de 100 à 500 copies (Pineau *et al.*, 2005). Parmi les cinq classes d'histones originellement caractérisées (H1, H2A, H2B, H3 et H4), nous avons décidé de créer nos différentes amorces à partir des quatre derniers et d'écarter dès le départ le gène H1, beaucoup moins conservé au sein de l'évolution (Wells et Brown, 1991; Kremer et Hennig, 1991). Parmi les quatre restant, H3 et H4 sont considérés comme les plus conservés. Entre autres, l'alignement des séquences protéiques codées par le gène H4 ne montre que deux substitutions d'acides aminés entre vertébrés et métaphytes (Bernhard et Schlegel, 1998) et le conservatisme du gène lui-même a permis de concevoir des amorces universelles capables d'amplifier le gène H4 chez tous les métazoaires (Pineau *et al.*, 2005). La divergence des séquences entre les plantes et les animaux chez H2A et H2B est à peu près dix fois plus grande que celle retrouvée chez H3 et H4 (Thatcher et Gorovsky, 1994), ce qui rendra plus difficile la conception d'amorces ciblant les premiers, vu l'absence de données de séquence provenant de génomes de mouches noires. Par ailleurs, Wilson et Post (1994) affirment que les variations retrouvées au niveau des espaceurs situés entre les gènes d'histones seraient potentiellement spécifiques de l'espèce, ce qui en ferait des cibles prometteuses, bien qu'à peu près complètement négligées jusqu'à maintenant.

Une fois les amorces de l'espaceur H3-H4 conçues et fabriquées, il a fallu établir les conditions optimales pour son amplification. Évidemment, les nombreuses variables liées au fonctionnement du PCR font en sorte qu'il est impossible de présumer qu'un seul type de conditions garantirait un succès dans toutes les situations. La longueur des amorces est directement liée à la complexité de l'ADN contenu dans l'échantillon et s'avère cruciale pour le succès de la PCR. Par ailleurs, la spécificité de la PCR dépend de la stabilité du complexe d'hybridation entre l'amorce et l'ADN cible. Différents facteurs s'associent pour déterminer cette stabilité, au premier plan desquels on compte la température de l'étape d'hybridation qui est principalement fonction de la température de fusion (T_m) de l'amorce. Guidée au départ par la valeur théorique du T_m , seule une approche empirique permettra ensuite de déterminer une température optimale ni trop haute, ni trop basse, ce dernier cas pouvant mener à une amplification non spécifique.

Étant donné que les conditions optimales ne peuvent pas être prévues à l'avance, nous avons donc procédé par tâtonnements afin de mettre au point les conditions pour l'amplification de chacune des cibles individuelles.

Une fois l'ADN amplifié, diverses techniques moléculaires seront appliquées afin de différencier les espèces. Dans un premier temps, nous avons établi la taille du produit PCR sur gel d'agarose et de polyacrylamide-urée. Advenant le cas où l'on se retrouvait avec des amplicons de tailles identiques ou quasi identiques, une caractérisation plus poussée au moyen du profil de restriction puis une évaluation du degré de divergence (technique des hétéroduplexes) seront effectuées.

L'utilisation des enzymes de restriction permet de partitionner l'ADN en multiples fragments de différentes tailles susceptibles d'être séparés par les techniques d'électrophorèse, établissant ainsi une empreinte typique à la fois de l'ADN digéré et de l'enzyme utilisé.

La technique des hétéroduplexes (fig 1.4) permet de détecter les mutations individuelles aussi bien que les divergences interspécifiques par la présence de mésappariements qui influencent la mobilité des molécules double brin dans des conditions particulières d'électrophorèse. Dans cette méthode, suite à une dénaturation et renaturation d'un mélange de fragments homologues d'ADN normaux et mutants (deux espèces distinctes dans notre cas), les hybrides moléculaires formés par l'association de ces deux chaînes d'ADN pourront dans certains cas produire un appariement imparfait. L'hybride ainsi formé possède sur un gel dénaturant d'acrylamide une mobilité réduite par rapport aux deux molécules originales parfaitement appariées (homoduplexe). Cette technique est largement empirique et ne permet pas de quantifier au sens strict le degré de divergence. Dans notre cas, elle permet surtout de mettre en évidence des différences non apparentes dans les profils de restriction. Elle a d'ailleurs permis ici de classer les espèces d'un même groupe.

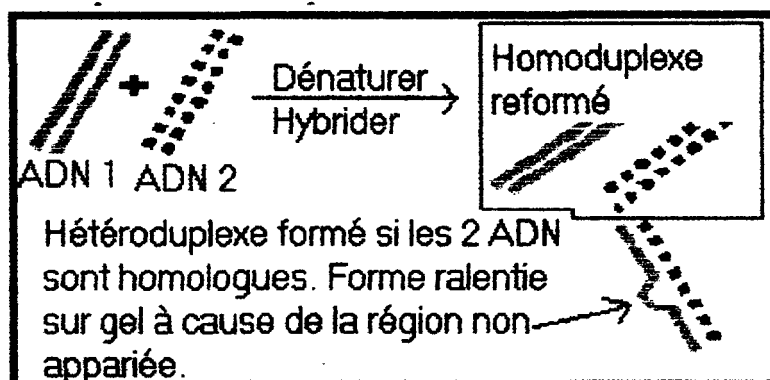


Figure 1.4 Schéma explicatif de la technique des hétéroduplexes.

RÉSULTATS

Étant donné que les résultats obtenus durant ma maîtrise font l'objet des deux articles présentés respectivement aux chapitres 2 et 3, je ne ferai qu'une brève revue des résultats ainsi obtenus. Le chapitre 2 rapporte les principaux résultats obtenus lors des essais concernant l'extraction d'ADN, tandis que nous retrouverons dans le chapitre 3 les résultats obtenus avec le premier marqueur (ITS-1) dans l'établissement de la clé d'identification moléculaire. Toutefois, avant ces chapitres, nous ferons aussi une brève évaluation des résultats préliminaires obtenus jusqu'à ce jour concernant les autres marqueurs utilisés.

La réussite d'un tel projet repose sur une méthode d'extraction d'ADN applicable de façon routinière tant en laboratoire que sur le terrain tout en fournissant du matériel de qualité PCR. Toutefois, même si l'ADN génomique extrait se présente souvent avant PCR en quantités trop petites pour être détectables sur gel d'agarose, il en subsiste assez pour fournir après amplification suffisamment du produit PCR requis pour la poursuite des expérimentations. De plus, même si la température de conservation des échantillons après extraction n'est pas critique, il reste que le tampon d'extraction, le mode de purification ainsi que le broyage de la larve nécessitent une attention particulière; si bien

que la meilleure combinaison est d'effectuer l'extraction en broyant la larve dans le tampon Bender et de purifier l'extrait sur résine Wizard (Promega).

Dans l'établissement d'une clé moléculaire de classification des mouches noires du Québec, nous avons évalué le potentiel de l'espaceur interne transcrit 1 (ITS-1) localisé entre les gènes d'ARNr 5.8S et 18S. Pour ce faire, l'ITS-1 de onze espèces de simulies, morphologiquement identifiées et réparties en trois genres, a été amplifié et les produits PCR obtenus ont été caractérisés tant par leur taille que par leurs profils de restriction issus de neuf endonucléases (Hae III, Alu I, Cfo I, Hinf I, Dde I, Rsa I, Msp I, Tru 9 I et Taq I). Puis, pour les espèces non différenciées par celles-ci, nous avons appliqué la technique des hétéroduplexes.

À elle seule, la taille des produits PCR sur gel d'agarose n'offre pas assez de pouvoir discriminant. Six des espèces faisant partie du genre *Simulium* ont une bande commune tandis que les autres se regroupent plutôt en trois niveaux différents et ce, tant sur gel d'agarose que de polyacrylamide. Toutefois, sur gel de polyacrylamide, nous observons pour la moitié des espèces la présence d'une seconde bande ayant une taille un peu plus élevée que celle retrouvée sur gel d'agarose, ce qui indiquerait la présence d'une hétérogénéité de longueur et de séquence chez l'ITS-1. Cette variabilité de taille des copies de l'espaceur à l'intérieur d'une même espèce pourrait selon nous résulter d'une région de type VNTR (répétitions en tandem en nombre variable). Ces courtes régions d'ADN présentent un motif répété et sont habituellement utilisées en génétique des populations comme marqueurs polymorphes intraspécifiques. Cette variabilité du nombre de répétitions s'expliquerait par le fait que lors de la réplication, l'ADN polymérase glisse sur ces motifs répétés et dérape en introduisant des erreurs de réplication. Ceci entraînerait une variation de longueur par addition ou délétion d'une ou de plusieurs unités de répétition. Le polymorphisme de longueur porterait donc sur la taille, dû aux nombreuses unités de répétitions. Les effets immédiats d'une mutation individuelle doivent être légers dans le cas de l'ADNr vu qu'il y a des centaines d'exemplaires de l'unité de transcription et les conséquences d'une mutation sont donc diluées au milieu du grand nombre de copies qui ont conservé leur type sauvage. Cela

s'applique d'autant plus à une séquence, comme celle de ITS-1, simplement destinée à l'élimination; les délétions/insertions présumées à l'origine de ce polymorphisme de taille subissent sans doute peu l'effet de la sélection naturelle et l'hétérogénéité observée reflète simplement la dérive génique tempérée par les mécanismes d'homogénéisation. Selon les concepts généralement admis (Lewin, 1999), la structure d'une famille multicopie résulterait de la balance entre les mutations individuelles créatrices de divergence entre copies et les mécanismes d'homogénéisation habituellement invoqués lors de processus d'évolution concertée, tels le « maître et l'esclave », la correction brusque et la recombinaison asymétrique. Il serait donc plausible que les organisations en tandem fournissent de fréquentes opportunités pour la formation de « mauvais appariement » entre des gènes dont les séquences sont les mêmes, mais qui se trouvent à des positions différentes dans leurs groupes. Ces « crossing-over » inégaux augmentent ou diminuent le nombre d'unités, ce qui rejoindrait l'idée de la variabilité de taille établie auparavant. Si une proportion appréciable de copies contient des délétions/insertions, cela se manifeste comme on le voit ici et tel que le rapportent Iranpour *et al.* (2005) par une multiplicité de bandes au niveau du produit PCR. Un pourcentage appréciable d'exemplaires mutants peut s'accumuler avant que leurs effets soient suffisamment importants pour que l'évolution les élimine. Certaines variantes pourraient même à la longue envahir l'ensemble du bloc d'unités tandem par recombinaison asymétrique et, via l'action combinée de la dérive génique et des mécanismes d'homogénéisation, remplacer la version antérieure.

Étant donné que la différence de taille entre les produits PCR ne peut distinguer les 12 espèces, nous avons donc effectué les profils de restriction. Bien que plusieurs enzymes n'ajoutent aucune information supplémentaire aux résultats obtenus précédemment, certains comme Alu I, Rsa I, Tru 9 I, Hinf I et Dde I, offrent un meilleur pouvoir de discrimination, mais sans pouvoir résoudre complètement les 12 espèces. Ce résultat quelque peu décevant contraste avec ceux d'Iranpour *et al.* (2005) qui a distingué avec succès entre 34 espèces de taons du Manitoba. Mais alors, pourquoi Iranpour *et al.* (2005) différencie-t-il si facilement ces 34 espèces alors qu'on a tant de difficulté à séparer nos 12 espèces de simulies? Cette faible différenciation d'espèces au sein des

simulies pourrait notamment être engendrée soit par une évolution ralentie de la famille ITS-1 chez les simulies néarctiques comparativement aux tabanides, soit à une origine plus récente des premières, en relation, selon l'hypothèse que nous présentons dans le chapitre 3, avec l'occupation par un faible nombre d'espèces fondatrices du territoire libéré par la dernière glaciation.

D'autre part, bien que quelques endonucléases nous permettent de différencier les espèces, il n'en demeure pas moins que nous remarquons le plus souvent la présence d'une ou plusieurs bandes communes à toutes, ce qui suggère qu'une partie de la séquence de l'ITS-1 serait fortement conservée. D'ailleurs cette question devra éventuellement être résolue par l'établissement de sa carte de restriction, sinon par le séquençage.

Le groupe des six espèces, du genre *Simulium*, qui ne sont résolues ni par la taille de ITS-1, ni par son profil de restriction, peut l'être par la technique des hétéroduplexes appliquée aux profils Dde I. Sur la base de la formation d'une structure inusitée apparentée à la jonction de Holliday formée lors de la recombinaison *in vivo* de l'ADN, toutes les espèces sont désormais différenciées à l'exception du groupe *S. venustum*/*S. decorum*. De plus, bien que cette technique fût aussi essayée sur le groupe *Prosimulium fuscum*/*fontanum*, celui-ci demeure encore non résolu. L'ensemble des données présentement disponibles nous permet donc de proposer un schéma de classification en deux étapes (fig 3.3).

Fait intéressant à remarquer, bien qu'aucun profil intermédiaire ne fût trouvé, les spécimens identifiés morphologiquement comme *S. decorum* présentent en réalité deux profils moléculaires bien différents. Le premier, *S. decorum**, ne produit sur gel d'acrylamide-urée qu'une seule bande, tandis que le second, *S. decorum*, produit une bande majeure et une mineure. Ce variant moléculaire représenterait donc un autre membre du complexe *S. decorum* (Adler et McCreadie, 1997). Évidemment, ce variant nous pousse à revoir la notion d'espèce et même à se poser la question suivante : mais qu'est-ce qu'une espèce? D'après le concept biologique de l'espèce (Harry, 2001), une

espèce est un groupe de populations naturelles réellement ou potentiellement interfertiles, mais en pratique isolées d'un point de vue reproductif des autres groupes. Trois critères utilisés seuls ou en association sont récurrents dans cette définition. Ce sont les critères morphologiques, de descendance et d'interfécondité/interstérilité. La seule prise en compte du critère morphologique pour la reconnaissance de l'espèce peut conduire à deux types d'erreurs, soit un regroupement au sein d'une même espèce de deux taxons de lignées évolutives différentes mais de morphologie identique, soit la reconnaissance de deux espèces différentes pour deux taxons de même lignée évolutive mais présentant une grande différence morphologique. Par contre, le critère de descendance donnerait à l'espèce sa cohérence et refléterait sa pérennité. La notion de descendance avec modification doit toutefois être introduite pour refléter l'évolution des espèces. Quant au critère d'interfécondité, il ne s'applique en situation naturelle qu'à des populations qui partagent la même aire de distribution. La définition proposée par Mayr est donc difficile à mettre en pratique et d'ailleurs, certains auteurs reprochent à cette définition de donner des indications sur le maintien du statut de l'espèce mais de ne pas donner une définition de ce statut (Harry, 2001).

Bien que le premier marqueur utilisé fournisse des résultats intéressants, il reste que celui-ci ne mène pas une distinction rapide, en une seule étape, de chacune des espèces à l'étude. Pour arriver à une telle méthode de diagnostic, seule l'utilisation de plusieurs marqueurs, appliqués soit dans des amplifications distinctes, soit simultanément au sein d'un PCR multiplexe, nous permettra d'atteindre notre but. Toutefois, peu importe le nombre de marqueurs utilisés, le design de plusieurs amorces (fixes et typiques d'espèce) est nécessaire ce qui requiert au niveau de la mise au point le recours à de multiples techniques (PCR, séquençage, clonage).

Étant donné l'ampleur du travail restant encore à réaliser et bien que quelques-unes des méthodes envisagées soient encore en développement, seuls l'approche appliquée pour le design des amorces et les résultats concernant la taille (sur gel d'agarose) de chacun des amplicons seront illustrés ici. Les 12 espèces analysées avec ces nouveaux marqueurs sont exactement les mêmes que pour l'ITS-1.

Les amorces amplifiant les gènes mitochondriaux proviennent directement de la littérature qui mentionne alors explicitement leur séquence (Tang *et al.*, 1998). La taille des produits d'amplification obtenus pour les marqueurs 12S, 16S et COII est non seulement similaire à celle rapportée par Tang pour le complexe *Psilopelmia*, mais aussi identique entre toutes les espèces présentes (Figs 1.5 à 1.7), et n'apporte donc aucune distinction supplémentaire par rapport à ITS-1. Vu la petite taille des amplicons obtenus, il apparaît également inutile d'en déterminer les profils de restriction car la probabilité de trouver des sites de coupure discriminants, même avec un endonucléase ayant une cible de 4 nucléotides, est relativement faible. De même, la caractérisation des hétéroduplexes obtenus à partir de l'espaceur entier semble peu prometteuse, puisqu'on s'attend, sauf surprise, à une distribution relativement compacte de la taille des hybrides formés. Cela nous ramène aux observations de Tang *et al.* (1998), qui n'a pu résoudre complètement le complexe *Psilopelmia* même en combinant les données de taille d'ITS-1 à l'analyse d'hétéroduplexes des marqueurs 12S, 16S et COII.



Figure 1.5 Produits PCR du gène mitochondrial COII : (a et n) standard Step Ladder 50 bp, (b) *S. longistylatum*, (c) *S. jenningsi*, (d) *S. aureum*, (e) *S. venustum*, (f) *P. fuscum*, (g) *S. decorum*, (h) *S. decorum**, (i) *S. vittatum*, (j) *P. fontanum*, (k) *C. mutata*, (l) *S. quebecense* et (m) *C. dacotensis*. Taille de l'amplicon : ~ 310 pb.

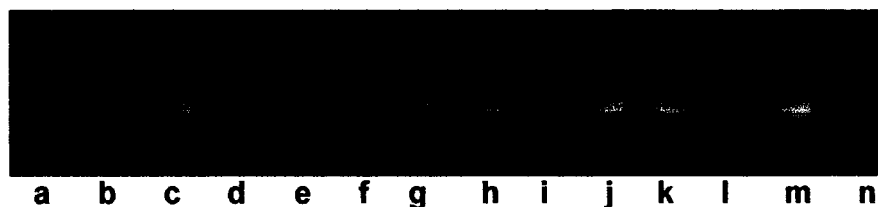


Figure 1.6 Produits PCR des gènes de l'ARNr mitochondrial 12S : (a et n) standard Step Ladder 50 bp, (b à m) mêmes espèces que pour la figure 1.5. Taille de l'amplicon : ~ 400 pb.

Les amorces ciblant l'espaceur entre les gènes d'histone H3 et H4 ont été conçues à partir des données de séquences de *D. melanogaster*, considérée comme espèce apparentée, et non à partir de celles, inexistantes, de similies. Cela complique un peu la tâche et requiert l'identification préalable des séquences des gènes de H3 et H4 ayant le plus de chance d'être identiques entre les deux. Nous avons donc comparé les séquences des gènes H3 et H4 entre *D. melanogaster* et *C. elegans*, espèce non apparentée servant de standard externe, dans le but de trouver les régions les plus conservées et en supposant qu'elles sont les mêmes chez les similies. Bref, la drosophile sert de banc d'essai pour la mise au point des amorces ciblant ces gènes. Cette région trouvée, il ne reste plus qu'à créer les différentes amorces. Comme mentionné plus haut, le choix des amorces est primordial pour la réussite du PCR et quelques critères, généraux ou spécifiques, doivent être respectés. Dans le cas des critères généraux, les amorces doivent comporter environ 50% de G + C et avoir préféablement des Tm identiques pour les deux amorces. Au niveau des critères spécifiques, on tient compte des différences possibles entre similie et drosophile dues à la dégénérescence du code génétique et on cherche à ancrer les amorces sur les régions où ces possibilités de dégénérescence sont réduites au minimum. Pour s'assurer d'une hybridation parfaite de l'extrémité 3' de l'amorce, celle-ci correspondra préféablement soit à un codon unique de méthionine (ATG) ou de tryptophane (TGG), soit à une séquence $G/C-G/C$ représentant les deux premières bases des codons d'un même acide aminé conservé. Quant aux codons internes à l'amorce, on introduira sur leur 3^e position une dégénérescence partielle (G/A ou C/T) ou totale, la désoxyinosine servant dans ce dernier cas de base indifférente capable d'accepter comme vis-à-vis n'importe quel partenaire.

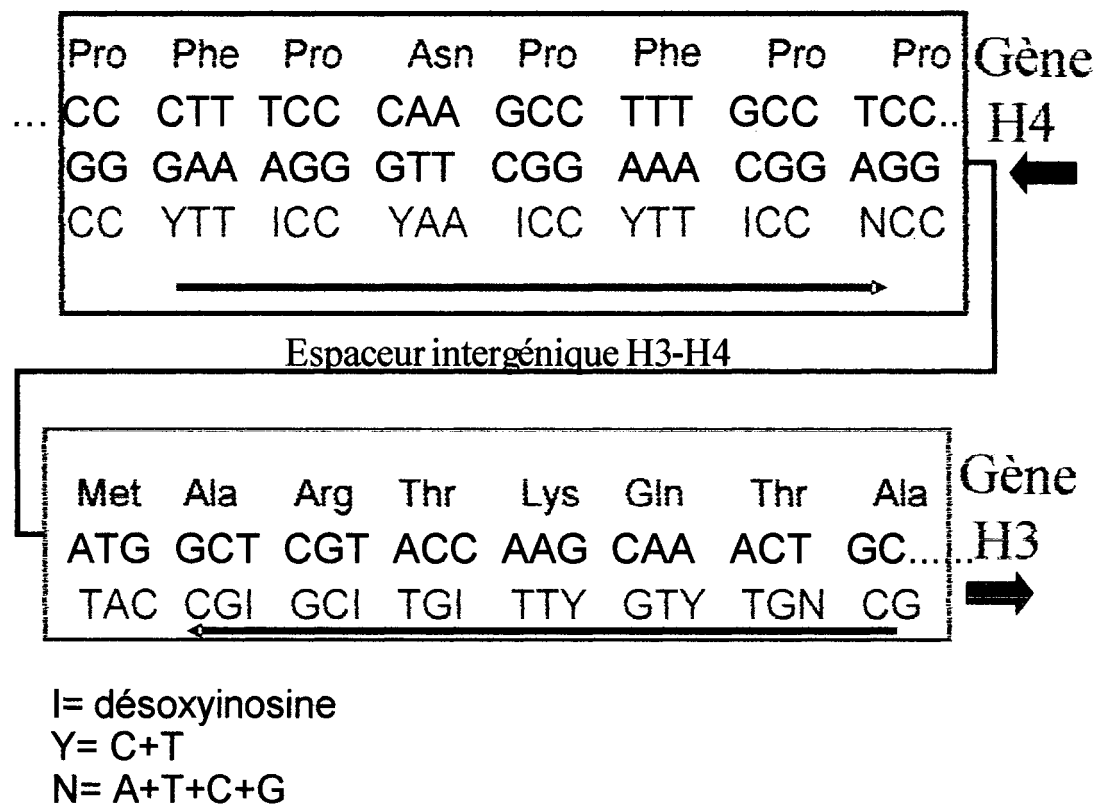


Figure 1.10 Séquences (en gris) des amorces ciblant l'espaceur des gènes des histones H3 et H4. Les flèches de couleur indiquent la direction de transcription du gène et la flèche grise l'orientation 5'→3' de l'amorce. En rouge : séquence d'acides aminés, avec la chaîne codante figurée en noir. En bleu : chaîne antisens.

Suite à la création de nos amorces, il fallait s'assurer de la validité de leur conception. Nous avons donc soumis l'ADN de *D. melanogaster* à quelques tests dans le but de vérifier si la taille de l'amplicon obtenu correspondait à celle prévue. De plus, ce même ADN a été utilisé pour optimiser les conditions d'amplification (rendement, spécificité) afin de rendre plus facile l'exécution d'un PCR chez les simules. Il est toutefois important de mentionner que des amorces conçues sur papier et comportant un niveau élevé de dégénérescence pourraient engendrer des problèmes subséquents quant à la présence en surplus de produits PCR non-spécifiques. Le contrôle effectué chez la drosophile (Fig 1.11) pour le marqueur H3-H4 donne un résultat positif et reproductible, malgré la différence de taille retrouvée entre les valeurs théorique (298 pb) et observée (362 pb). Bien que cette différence puisse sembler assez grande, nous avons toutefois

observé qu'il n'existe pas une uniformité absolue au sein de la littérature et des banques de données. On y retrouve souvent des différences pouvant atteindre 100 pb, ce qui peut ne pas surprendre dans le cas d'une famille de séquences répétées dont on aurait étudié des membres différents. Par ailleurs et contrairement aux marqueurs mitochondriaux, on obtient chez les similies une diversité de taille du produit PCR, qui se distribue en approximativement quatre niveaux sur le gel d'agarose. Dans le cas de *S. venustum* et *S. decorum*, la bande obtenue est clairement double, indiquant comme pour ITS-1 une hétérogénéité de taille et/ou de séquence. L'espaceur H3-H4 ne permet toujours pas une distinction radicale des espèces, mais l'emplacement de celles-ci à l'intérieur d'un schème de classification diffère légèrement de celui qui caractérise ITS-1, de sorte que la combinaison des deux séries d'information mène à une classification plus performante que si on considère séparément les deux marqueurs; en réalisant un PCR multiplexe, on obtiendrait théoriquement le diagramme de la figure 1.12. Donc, en une seule étape et sans utiliser d'enzyme de restriction, on arrive à une différenciation des espèces égale, sinon supérieure à celle obtenue avec ITS-1 digéré par Dde I ou Alu I. Éventuellement, l'ajout d'un 3^e, puis d'un 4^e marqueur, pourrait mener à l'élaboration d'un code-barre absolument distinctif pour chaque espèce (Fig 1.13).

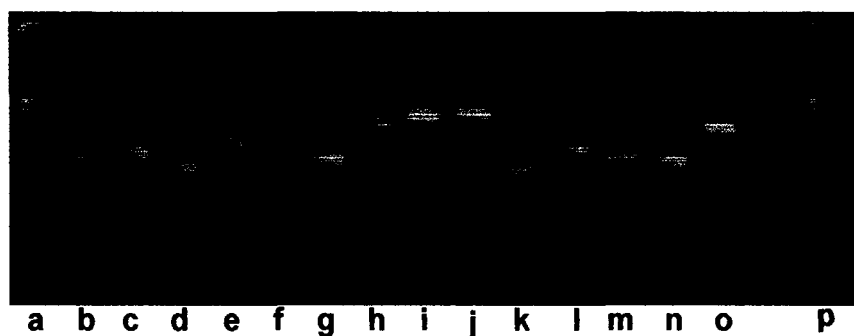


Figure 1.11 Taille de l'espaceur H3-H4 : (a et p) standard pUC18-HaeIII, (b et c) *D. melanogaster* (d) *S. jenningsi*, (e) *S. aureum*, (f) *S. venustum*, (g) *P. fuscum*, (h) *S. decorum*, (i) *S. decorum**, (j) *S. vittatum*, (k) *P. fontanum*, (l) *C. mutata*, (m) *S. quebecense*, (n) *C. dacotensis* et (o) *S. longistylatum*. Les conditions d'amplification sont similaires à celles utilisées pour ITS-1 (chapitre 3), sauf pour les concentrations d'amorces qui sont triplées et pour une température d'hybridation augmentée à 52°C.

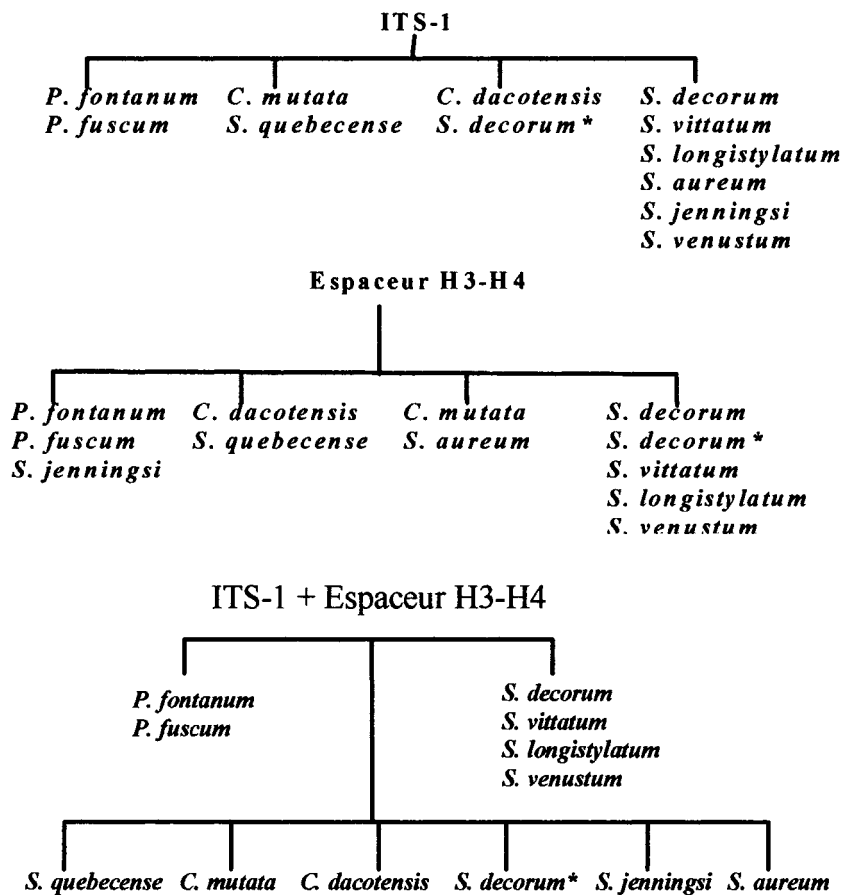


Figure 1.12 Schème de classification de 12 espèces de mouches noires pour les marqueurs ITS-1 et H3-H4 seuls ou combinés.

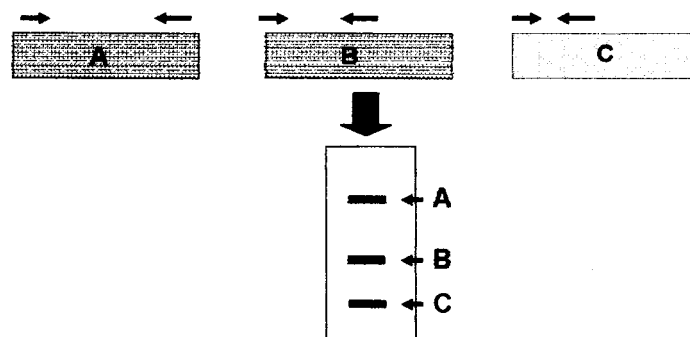


Figure 1.13 Principe du PCR multiplexe. Amplification simultanée de différentes cibles A, B, C à l'aide d'un mélange de plusieurs paires d'amorces.

Par ailleurs, comme les amorces de l'espaceur H3-H4 ont été créées à partir de séquences provenant de banques de données et que celles-ci fournissent également la séquence complète du bloc comprenant les gènes des cinq types d'histones, rien ne nous empêche donc en principe d'agrandir la région ciblée soit de H2 (A ou B) à H3 et d'en mettre individuellement en évidence chacun des espaceurs dans le cadre d'un PCR multiplexe. Alternativement, on pourrait amplifier le bloc en totalité, puis, comme pour ITS-1, en digérer ensuite le produit PCR de façon à augmenter, grâce à des profils de restriction discriminants, la quantité d'information disponible.

Conclusion

Bien que plusieurs résultats aient été accumulés durant ces deux dernières années, il reste encore pas mal de chemin à parcourir avant de disposer d'une méthode moléculaire universelle et simple de classification. Ces travaux se poursuivront au doctorat, où on prévoit non seulement améliorer les méthodes de typage moléculaire, mais ajouter nombre de sorties de terrain afin d'élargir l'éventail des espèces étudiées et d'ainsi mieux valider l'universalité d'une clé moléculaire d'abord élaborée à partir d'un groupe restreint de 12 espèces seulement. Cependant, il demeure possible qu'avec un nombre très élevé d'espèces, on ne puisse toutes les distinguer même en mode de PCR multiplexe et qu'on obtienne parfois des groupes non résolus. D'ailleurs, ce problème se pose déjà dans le cadre de la classification morphologique lorsqu'on tombe sur des complexes d'espèces-sœurs et qu'on doit alors faire appel à la cytogénétique. Dans ce type de situation, l'approche moléculaire devra sans doute être modifiée, quitte à aller jusqu'au clonage et au séquençage des marqueurs étudiés. À partir de là, on pourrait utiliser les données de séquence pour mettre au point un PCR sélectif exploitant les différences individuelles dues aux mutations ponctuelles ou étendues dans le but de concevoir toute une batterie d'amorces dont chacune serait sélective d'une seule espèce (Fig 1.14).

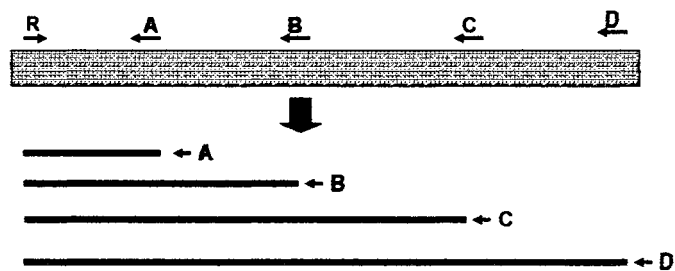


Figure 1.14 Principe du PCR sélectif de l'espèce. La technique utilise à une extrémité de la cible une amorce commune R ancrée sur une région conservée et à l'autre une série d'amorces spécifiques A, B, C, D dont chacune cible une séquence absolument particulière à chaque espèce. Les sites d'hybridation des amorces sont prévus de façon à générer des produits PCR de tailles distinctes.

BIBLIOGRAPHIE

- Adler, P.H., Currie, D.C. et Wood, D.M., 2004. The blackflies (Simuliidae) of North America. ROM publication in sciences. NewYork, USA.
- Adler, P.H. et Mc Creadie, J.W., 1997. The hidden ecology of black flies: sibling species and ecological scale. American Entomologist. Fall: 153-161.
- Bernhard, D. et Schlegel, M., 1998. Evolution of histone H4 and H3 genes in different ciliate lineages. Journal of Molecular Evolution. 46: 344-354.
- Douglas C.A., 2004. Three new species of *Inseliellum* (Diptera: Simuliidae) from Polynesia. Zootaxa. 450: 1-18.
- Harry, M. 2001. Génétique moléculaire et évolutive. Édition Maloine. Paris.
- Hwang, Ui-Wook et Kim, Won, 1999. General properties and phylogenetic utilities of nuclear ribosomal DNA and mitochondrial DNA commonly used in molecular systematics. The Korean Journal of Parasitology. 37: 215-228.
- Iranpour, M., Schurko, A.M., Klassen, G.R. et Galloway, T.D., 2004. DNA fingerprinting of tabanids (Diptera: Tabanidae) and their respective egg masses using PCR-restriction fragment profiling. The Canadian Entomologist. 136: 605-619.
- Kremer, H. et Hennig, W., 1991. Isolation and characterization of a *Drosophila hydei* histone DNA repeat unit. Nucleic Acids Research. 18: 1573-1580.
- Krüger, A., Gelhaus, A. et Garms, R., 2000. Molecular identification and phylogeny of East African *Simulium damnosum* S.l. and their relationship with West African species of the complex (Diptera: Simuliidae). Insect Molecular Biology. 9: 101-108.
- Lewin, B. 1999. Gènes VI. De Boeck Université. Paris.
- Mank, R., Wilson, M.D., Rubio, J.M. et Post, R.J., 2004. A molecular marker for the identification of *Simulium squamosum* (Diptera: Simuliidae). Annals of Tropical Medecine & Parasitology. 98:197-208.
- Morales-Hojas, R., Post, R.J., Cheke, R.A. et Wilson M.D., 2002. Assessment of rDNA IGS as a molecular marker in the *Simulium damnosum* complex. Medical and Veterinary Entomology. 16: 395-403.
- Morlais, I. et Severson, D.W., 2002. Complete mitochondrial DNA sequence and amino acid analysis of the cytochrome C oxidase subunit I (COI) from *Aedes aegypti*. DNA Sequence. 13: 123-127.
- Pineau, P., Henry, M., Suspène, R., Marchio, A., Dettai, A., Debruyne, R., Petit, T., Lécu, A., Moisson, P., Dejean, A., Wain-Hobson, S. et Vartanian, J-P., 2005. A universal primer set for PCR amplification of nuclear histone H4 genes from all animal species. Molecular Biology and Evolution. 22: 582-588.

- Post, R.J., Flook, P.K. et Millest, A.L., 1993. Methods for preservation of insects for DNA studies. *Biochemical Systematics and Ecology*. 21: 85-92.
- Rooney, A.P., Piontkivska, H. et Nei, M., 2002. Molecular evolution of the nontandemly repeated genes of the histone 3 multigene family. *Molecular Biology and Evolution*. 19: 68-75.
- Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H. et Flook, P., 1994. Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America*. 87: 651-701.
- Steiner, J.J., Poklemba, C.J., Fjellstrom, R.G. et Elliott, L.F., 1995. A rapid one-tube genomic DNA extraction process for PCR and RAPD analyses. *Nucleic Acids Research*. 23: 2569-2570.
- Szalanski, A.L. et Owens, C.B., 2003. Sequence change and phylogenetic signal in muscoid COII DNA sequences. *DNA Sequence*. 14: 331-334.
- Tang, J., Toè, L., Back, C. et Unnasch, T.R., 1995a. Mitochondrial alleles of *Simulium damnosum* sensu lato infected with *Onchocerca volvulus*. *International Journal of Parasitology*. 25: 1251-1254.
- Tang, J., Toè, L., Back, C., Zimmerman, P.A., Pruess, K., et Unnasch, T.R., 1995b. The *Simulium damnosum* species complex: phylogenetic analysis and molecular identification based upon mitochondrially encoded gene sequences. *Insect Molecular Biology*. 4: 79-88.
- Tang, J., Pruess, K. et Unnasch T.R., 1996. Genotyping North American black flies by means of mitochondrial ribosomal RNA sequences. *Canadian Journal of Zoology*. 74: 39-45.
- Tang, J., Moulton, J.K., Pruess, K., Cupp, E.W. et Unnasch, T.R., 1998. Genetic variation in North American blackflies in the subgenus *Psilopelmia* (*Simulium*: Diptera: Simuliidae). *Canadian Journal of Zoology*. 76: 205-211.
- Thatcher, T.H. et Gorovsky, M.A., 1994. Phylogenetic analysis of the core histones H2A, H2B, H3 and H4. *Nucleic Acids Research*. 22: 174-179.
- Wells, D. et Brown, D., 1991. Histone and histone gene compilation and alignment update. *Nucleic Acids Research*. 19: 2173-2188.
- Wilson, M.D. et Post, R.J., 1994. Integration of morphometric, cytogenetic and molecular techniques: a case study of *Simulium damnosum*. Dans «Identification and characterization of pest organisms», ed. Hawksworth, D.L. pp. 215-224. Wallingford, U.K.: CAB International.
- Wood, D.M., Peterson, B.V., Davies, D.M. et Gyorkos, H., 1963. The Black flies (Diptera: Simuliidae) of Ontario. Part II. Larval identification, with description and illustration. *Proceedings of the Entomological Society of Ontario*. 93: 99-129.
- Xiong, B. et Kocher T.D., 1991. Comparison of mitochondrial DNA sequences of seven morphospecies of black flies (Diptera: Simuliidae). *Genome*. 34: 306-311.

CHAPITRE II

Note scientifique

PCR suitability of field-extracted DNA from small invertebrates

Christine Gaudreau¹
Bernard LaRue¹
Guy Charpentier^{1*}

¹Département de Chimie-biologie, Université du Québec à Trois-Rivières,
3351 Boul. des Forges, Trois-Rivières, Qc., Canada G9A5H7

* **Note.** Author to whom all correspondence should be sent : Guy Charpentier, Département de Chimie-biologie, Université du Québec à Trois-Rivières, Trois-Rivières, Canada G9A5H7. Phone : 1-819-376-5053. FAX : 1-819-376-5084. E-Mail : guy.charpentier@uqtr.ca

Key words : Simuliidae, black flies, DNA field extraction.

ABSTRACT

A low technology and field friendly method to extract DNA from small invertebrates is described. Extracted DNA can be stored for at least ten days at ambient temperature, without loss of amplification ability, before final resin-based purification is performed in the laboratory. Using this method, one can perform up to 4000 PCR reactions with DNA purified from a single black fly larva.

RÉSUMÉ

On présente une méthode d'extraction d'ADN de petits invertébrés, qui demeure simple et applicable lors de travaux sur le terrain. Le matériel extrait peut être conservé, sans perte de capacité d'amplification, pour au moins dix jours à température ambiante avant l'étape finale de nettoyage sur résine de silice. L'ADN extrait d'une seule larve de mouche noire fournit assez de matériel pour effectuer 4000 réactions PCR.

Published methods for DNA purification, probably reaching into the hundreds, vary not only according to the material being extracted, but also differ by the required amount of work, by their complexity and by their cost. In recent years, commercial kits based on porous silica resins have largely replaced the previous and more tedious manual methods, but these kits can hardly be used when laboratory facilities remain minimal, as occurs during prolonged field trips.

We recently undertook a study in which Simuliidae (black flies) larvae and their nematode parasites were collected at several remote stations. These organisms are of small size and yield only limited amounts of DNA: for instance the dry weight of a 4th instar black fly larva, most of it cuticle, averages 1 mg and a fresh larva contains about 160 ng of extractable DNA. During field work, the samples should preferably be processed for crude DNA extraction as soon as possible, with the final purification taking place later in the laboratory. Under those circumstances, several requirements should be met: stability of the extraction buffer, simplicity of the extraction method, stability of the extracted DNA for at least several days at ambient temperature and finally the ability to operate under conditions where cross-contamination of samples is prevented, especially if PCR typing has to be performed later on.

With these considerations in mind, we first sterilized in the laboratory, packs of pipette tips and of 1.5 mL microcentrifuge tubes with their plastic pestles and prepared two extraction buffers according to already published recipes. Bender's buffer (1) contains 100 mM Tris-HCl, pH 9.1, 100 mM NaCl, 200 mM sucrose, 50 mM EDTA, 0.5% sodium lauryl sulphate and 0.8 mg/mL proteinase K (ProK). The so-called ROSE buffer of Steiner *et al.* (2) is 10 mM Tris-HCl, 312 mM EDTA, pH 8.0, 1% sodium lauroyl sarcosine and 1% polyvinylpolypyrrolidone. In the field, samples are transferred to microcentrifuge tubes using flame-sterilized forceps and 125 μ L of either extraction buffer is then added. After grinding the insect for 10-20 seconds, the tube is capped and then floated in a small portable water-bath at approximately 65°C (Bender) or 90°C (ROSE) for 60 or 20 minutes respectively; temperature is maintained by the use of a small portable oven or alcohol burner. Exact timing or temperature does not seem to be crucial for the extraction to be successful. Back in the laboratory, 1/10 volume of 8 M CH₃COOK is added to samples extracted with Bender's buffer, which are then kept on ice for 20 minutes. Finally, samples are centrifuged and their supernatant is mixed with 0.5 mL of Wizard DNA clean-up resin (Promega) and purified according to manufacturer's specifications. Alternatively, ProK was omitted during Bender extraction, or the extracted DNA was ethanol precipitated (EtOH) instead of resin cleaning. In all cases, DNA is recovered in 50 μ L of sterile water of which 10 μ L is analyzed on 1% agarose gel.

As shown in Fig. 2.1, crude extracts of the black fly *Cnephia mutata* are characterized by a prominent smear, most likely RNA fragments, ahead of the gel. High molecular weight DNA from individual larvae, expected at the same level as the salmon DNA standard, is not always visible, which is not too surprising since the amounts being analyzed are close to the limit of detection. The RNA smear disappears from resin-cleaned samples, with again the upper DNA band not always detected. With Bender

extraction, attempts were also made to omit the grinding step in the hope that the high temperature and detergent concentration used for extraction would allow DNA to leak out of the cuticle into the surrounding medium. However, in crude samples, the usual RNA smear was not visible even if the specimen is soaked for 16-18 hours at 65°C, which suggests that no nucleic acid was recovered.

As mentioned above, the absence of a visible DNA band does not preclude the success of PCR typing. For this purpose, we amplified the intergenic spacer between the 18S and 5.8S nuclear RNA genes (ITS-1) according to Tang *et al.* (3), with the appropriate dilution of extracted DNA amounting to one fifth of the total reaction volume. At first, inconsistent results were obtained from samples which did not undergo the grinding step, with only a few of them producing a successful amplification and for definite dilutions only (data not shown). On the other hand and up to a 200-fold dilution, all resin-cleaned DNA samples, irrespective of the extraction buffer being used, yielded an obvious PCR product having the expected size for *C. mutata* ITS-1 (Fig. 2.2). At this dilution (1/200), it was calculated that a single black fly larva produces enough material for about 4000 amplification reactions. Contrary to Steiner *et al.* (2), a simple 200-fold dilution of crude ROSE extracted DNA prior to amplification does not yield any PCR product. This can likely be attributed to the very high EDTA content of ROSE buffer (312 mM), which leads to Taq DNA polymerase inhibition unless extreme dilutions are used; with the type of sample analyzed here, this dilution would simply mean the loss of the PCR signal. A 10-fold reduction of EDTA concentration at the extraction step did not improve this result (data not shown). In the case of Bender-extracted material, ethanol precipitation can substitute for resin cleaning, which represents a cheaper alternative to silica resins and mini-columns. However, one then notices consistent amplification only for dilution factors of 20 or more, which suggests the removal through dilution of impurities inhibiting Taq polymerase. DNA recovered by ethanol precipitation could probably not be used for RAPDs, since this procedure requires 10-20 ng of DNA for clear-cut results when applied to insect DNA (1). Moreover, the inhibition of ITS-1 amplification seems more pronounced if Bender's buffer contains ProK. Since the latter does not seem to improve DNA extraction as such, it should preferably be omitted from the procedure.

Finally, crude extracted DNA was brought back to the lab and stored in the dark at various temperatures, namely -20, 4 and 25°C, for ten days prior to resin cleaning and amplification. As shown by the representative data of Fig. 2.3, all samples gave the expected PCR product, with no obvious difference related to storage temperature. Our method can be applied during extended field trips performed under low technology conditions. It is likely that the storage time at 25°C could be extended somewhat without loss of amplification capacity. At best, one notices that DNA extracted using ROSE buffer yields a slightly weaker signal, which could be related to the integrity and to the average molecular weight of the recovered DNA. The current procedure was also successfully applied to a variety of insects, including other diptera, coleoptera and lepidoptera and to mermithids, that are nematode parasites of black fly larvae. In summary, grinding followed by DNA extraction using Bender, but omitting ProK, seems the best recipe for obtaining PCR-quality DNA from small invertebrates.

REFERENCES

1. L  ry, X., LaRue, B., Cossette, J. and Charpentier, G. (2003) Characterization and authentication of insect cell lines rising RAPD markers. *Insect Biochem. Mol. Biol.* **33**, 1035-1041.
2. Steiner, J.J., Poklemba, C.J., Fjellstrom, R.G. and Elliot, L.F. (1995) A rapid one tube genomic DNA extraction process for PCR and RAPD analyses. *Nucleic Acids Res.*, **23**, 2569-2570.
3. Tang, J., Moulton, J.K., Pruess, K., Cupp, E.W. and Unnasch, T.R. (1998) Genetic variation in North American black flies in the subgenus *Psilopelmia* (Simulium: Diptera: Simuliidae). *Can. J. Zool.*, **76**, 205-211.

Figure 2.1: Agarose gel electrophoresis of black fly larva DNA extracted by various methods. St, salmon DNA standard, 50, 100, 150 and 200 ng respectively; L, λ /*HindIII* + *EcoRI* molecular weight marker; 1 and 2, ROSE (312 mM EDTA); 3 and 4, ROSE (312 mM EDTA) followed by resin cleaning; 5 and 6, ROSE (31.2 mM EDTA); 7 and 8, Bender without and with ProK respectively, followed by EtOH; 9 and 10, soaked (no grinding) in Bender without and with ProK, followed by EtOH; 11 and 12, Bender + resin cleaning, without and with grinding respectively; 13, pool of five larvae purified by Bender + resin.

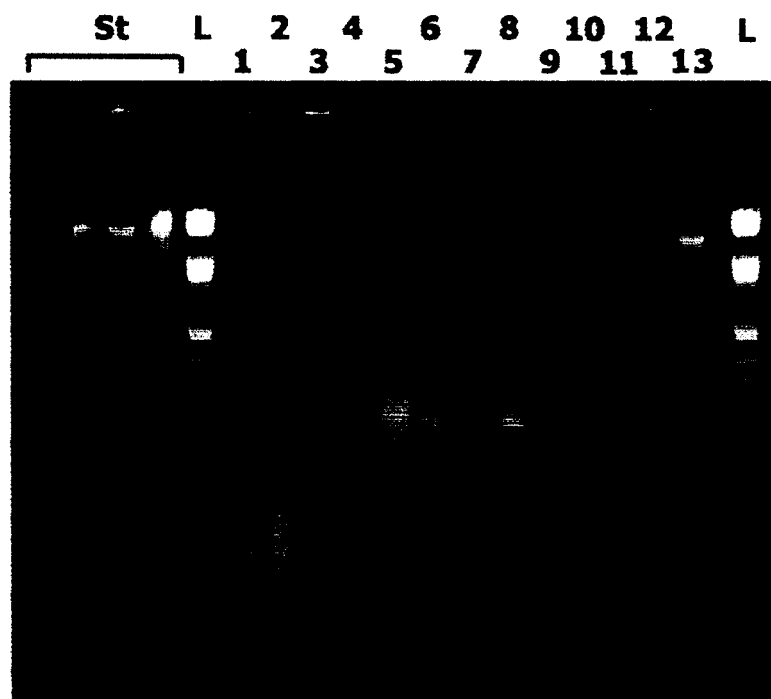


Figure 2.2: Effect of dilution of black fly DNA on the PCR signal of ITS-1. Upper panel: post-extraction resin cleaned DNA. Lower panel: ethanol precipitated DNA after Bender extraction.

From left to right: no dilution, 5, 20, 50, 100 and 200-fold diluted. L: λ HindII + EcoRI molecular weight marker.

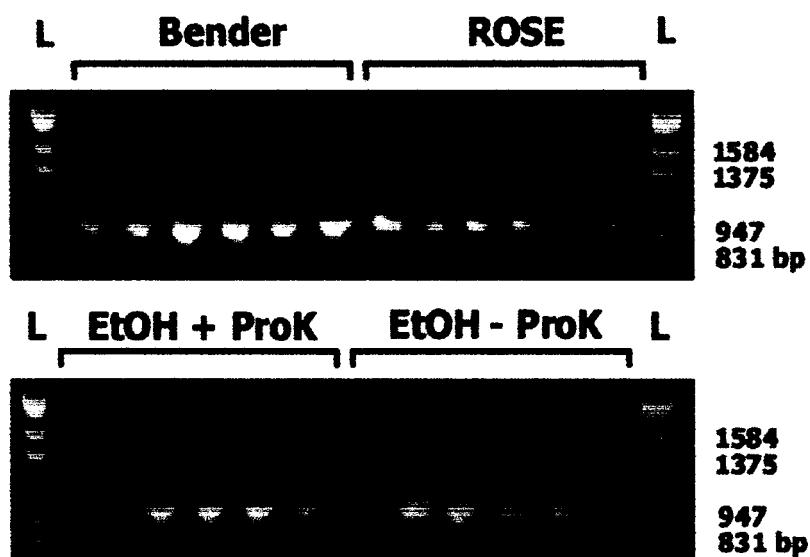
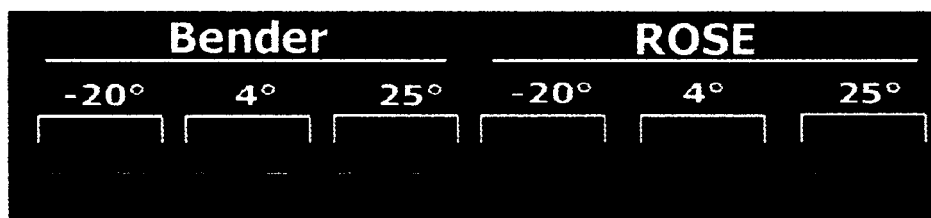


Figure 2.3: Effect of storage temperature of field-extracted DNA on the PCR signal. Samples were kept for ten days at the indicated temperature preceding resin cleaning and diluted 20-fold before amplification.



CHAPITRE III

Article scientifique

**Characterization of black flies (Diptera: Simuliidae) from
Quebec (Canada) by restriction-PCR**

**Christine Gaudreau¹
Bernard LaRue¹
Guy Charpentier^{1*}**

¹Département de Chimie-biologie, Université du Québec à Trois-Rivières,
3351 Boul. des Forges, Trois-Rivières, Qc., Canada G9A5H7

* **Note.** Author to whom all correspondence should be sent : Guy Charpentier, Département de Chimie-biologie, Université du Québec à Trois-Rivières, Trois-Rivières, Canada G9A5H7. Phone : 1-819-376-5053. FAX : 1-819-376-5084. E-Mail : guy.charpentier@uqtr.ca

Key words : Simuliidae, black flies, ITS-1, restriction-PCR, HDA

Abstract

The ITS-1 spacer located between the 5.8S and 18S rRNA genes from 12 black fly species from Quebec was amplified and PCR products were characterized first by their size and next through fingerprints obtained with nine different restriction enzymes each having a 4 base-pair target sequence. Spacer length alone leads to a classification based on four distinct groupings. In the case of unresolved groups, it was found that the Dde I profile combined with heteroduplex analysis (HDA) fragments could resolve all 12 species, with the exception of two groups, namely *P. fontanum/fuscum* and *S. venustum/decorum*. Furthermore, specimens initially classified as *S. decorum* according to morphology split into two different DNA profiles, likely members of a complex of siblings. Although based on a limited number of species, this work suggests that ITS-1 typing could constitute, with other markers to be added later on, the first step in establishing a universal molecular key for black flies.

Résumé

Les produits PCR obtenus par amplification de l'espaceur ITS-1 localisé entre les gènes nucléaires d'ARNr 5.8S et 18S de 12 espèces de mouche noire du Québec ont été caractérisés tant par leur taille que par leurs patrons de restriction, issus de neuf endonucléases possédant chacune une séquence cible de 4 pb. Alors que la taille de l'espaceur ne réussit à subdiviser cet échantillonnage qu'en quatre groupes distincts, les profils de restriction Dde I combinés lorsque nécessaire à l'analyse des hétéroduplexes permettent de résoudre les 12 espèces les unes des autres, à l'exception des couples *P. fontanum/fusum* et *S. venustum/decorum*. De plus, les spécimens originellement tous classifiés comme *S. decorum* selon les critères morphologiques fournissent en fait, sans profil intermédiaire, deux variants moléculaires représentant sans doute deux membres d'un complexe d'espèces. Malgré le nombre limité d'espèces étudiées, ce travail suggère que ITS-1 pourrait constituer le premier d'une série de marqueurs génétiques à employer dans l'établissement d'une clé moléculaire universelle des mouches noires.

Introduction

Black flies are medically important as vectors of *Onchocerca volvulus*. Besides disease transmission, they constitute a severe nuisance for humans and are responsible for economic losses in livestock, poultry and wildlife (Crosskey 1990). For instance, *Simulium arcticum* and *S. luggeri* are agricultural pests in Canada (Fredeen 1977, 1985) and *S. vittatum* preys on large domestic and wild mammals (Rothfels and Featherston 1981).

Classification of insect species is critical for both basic and applied research. However, black flies that belong to the more primitive nematoceran Diptera, present little morphological diversity (Xiong and Kocher 1993). This problem arises in many groups because of small insect size, environmentally-dependent morphological attributes and the prevalence of biotypes and complexes of species (Shouche and Patole 2000). Based on fixed-inversion differences in polytene salivary gland chromosomes of larvae, many nominal species now appear as complexes of siblings (Zhu *et al.* 1998). However, karyotyping does not allow the identification of adult female flies (Krüger *et al.* 2000). This represents a major problem, since the ability to identify all stages (Carew *et al.* 2003), principally adult, of insect vector species matters most in helping to understand their different roles in pathogen transmission dynamics and in reducing disease incidence through selective vector control (Morales-Hojas *et al.* 2002). More recently, DNA amplification methods have been added to the morphological and cytogenetic tools. In contrast to traditional keys, molecular methods may be applied equally well to

both sexes and to all stages of the insect life cycle, thus abolishing the need for multiple classification keys.

Many PCR targets seem useful to identify insects, particularly Diptera. Mitochondrial (mt) DNA is widely used for this purpose (Miller *et al.* 1999, Muraji and Nakahara 2002, Salazar *et al.* 2002). Tang *et al.* (1996, 1998) proposed to use mitochondrial genotyping based on heteroduplex analysis of amplification products derived from the cytochrome oxidase II (COII), 12S and 16S genes. However, according to Xiong and Kocher (1993), the 16S rRNA gene does not constitute a good taxonomic marker among closely related taxa, since only a relatively small proportion of sites varies freely, which leaves few characters available for phylogenetic analysis. Moreover, some inconsistencies show up when different parts of the 16S rRNA are used to infer the phylogeny. With the mt cytochrome oxidase subunit 1 gene, Linton *et al.* (2002) succeeded in separating five members of the *Culicoides imicola* species complex. According to Pruess *et al.* (2000), the COII gene investigated at the sequence level seems to be useful at the subgeneric level, but much less at higher taxonomic levels. Also, random amplified polymorphic DNA (RAPD) distinguishes between sibling species of some Diptera and aphids (Ballinger-Crabtree *et al.* 1992, Sebastiani *et al.* 2001). RAPD band patterns have been used to determine the degree and pattern of genetic differentiation in insect populations (de Sousa *et al.* 1999, Manguin *et al.* 1999, Ross *et al.* 1999) and to authenticate the origin of insect cell lines (Léry *et al.* 2003).

Most of these works attempt to discriminate between siblings or deal with population genetics, but, to our knowledge, little effort has been directed at the establishment of a

universal molecular key for black flies. Requirements for developing a key would be: a) to find several DNA markers, each of which can be amplified with universal primers, b) to show the constancy of the DNA fingerprint within each species, c) to evaluate the ability of these markers, first individually and collectively, to identify distinct species and d) to validate the key through correlations with morphological taxonomy. This would exclude the use of RAPD or minisatellites, which are best at revealing differences between populations or individuals of the same species, but involve such a high degree of intraspecific polymorphism that they cannot be routinely applied to molecular taxonomy. On the other hand, the internal transcribed spacer (ITS-1) between 5.8S and 18S rDNA is relatively free of structural and functional constraints, which allows it to evolve quite rapidly (Tang *et al.* 1996). rDNA is one of the favourite marker in molecular systematics (Brower and DeSalle 1994, Caterino *et al.* 2000), because its repetitive structure allows for easy detection. Moreover, sequence divergence between individual copies remains generally low, which explains its use as a model for the study of molecular drive (Strachan *et al.* 1985, Dover 1986) and concerted evolution (Arnheim 1983) in repeated-sequence families (Tang *et al.* 1996). Previous work on the 18S-28S ITS from the *S. damnosum s.l.* complex has demonstrated that size polymorphism detectable by PCR amplification of the entire ITS can distinguish at least two of the sibling species of this complex (Brockhouse *et al.* 1993).

In black flies, ITS-1 size alone lacks discriminating power, since, in most cases, it varies narrowly between 800 and 1300 base pairs and is not directly related to its sequence itself. At the other extreme, the complete ITS-1 nucleotide sequence could well lead to a very accurate classification, but this hardly appears feasible for routine applications. In

addition, this approach neglects the problems arising from the possible sequence variants of a substantial fraction of the rDNA family copies. As a midrange compromise, restriction fragment patterns offer partial insights of the sequence itself (Avisé *et al.* 1987, Satta *et al.* 1988, Iranpour *et al.* 2004), and can be obtained easily for a large number of specimens and should not be significantly influenced by minor length and sequence variants. Due to the relatively small size, to the paucity of sequence data from black fly ITS-1, and since the position of the restriction sites are unknown, we tested a diversity of arbitrarily chosen restriction endonucleases with short (4-bp) specific target sequences. On a random sequence basis, each enzyme should, on the average, generate three or four DNA fragments, some of them, hopefully, giving far more information than ITS-1 size alone. It remains to be determine which enzymes or combination of them will provide a distinctive signature for as many species as possible. For this purpose, we randomly collected a group of 11 morphological (later on rising to 12 according to molecular criteria, see below) species from southeastern Canada.

Materials and methods

Collection of black flies

Black flies used in this study were collected as larvae from February to June 2003 and during May 2004 from numerous streams in the Saint-Maurice faunal reserve, at St-Élie, at Lac Souris and in the Ottawa region. The collection points are summarized in Table 1. Black fly larvae were fixed in 95% ethanol, which is optimal for routine DNA preservation (Post *et al.* 1993), and kept on ice until storage at 4°C in the laboratory. They were later identified using a binocular according to morphology keys from Wood *et al.* (1963) and Adler *et al.* (2004).

DNA extraction

After five minutes of rehydration at room temperature in TE buffer (Tris-HCl 0.01 M pH 8.0, EDTA 0.001 M), larvae were homogenized and DNA was extracted and purified on a commercial resin as described elsewhere (Léry *et al.* 2003). Purified DNA was eluted from the resin in 50 µL of sterile water and stored at -20 °C. One larva yielded usually between 50 and 200 ng of DNA.

DNA amplification

PCR was carried out in a total volume of 25 µL in a solution containing 60 mM Tris-HCl (pH 8.3), 15 mM (NH₄)₂SO₄, 2.5 mM MgCl₂, 200 µM of each dNTP, 10 pmol of each primer, 1 unit of Taq DNA Polymerase (Roche) and 5 µL of a 1/20 dilution of template DNA, prepared as described above. The cycling conditions were 3 min at 95 °C + 38 cycles (30s at 95 °C, 50s at 46.5 °C, 60s at 72 °C) + 2 min at 72 °C. Primers for

ITS-1, as described by Tang *et al.* (1998) were: forward 5' TGGTGCATGGCGTTCTTAG 3', reverse 5' GTCGATGTTTCATGTGTCCTGC 3'. All PCR products were characterized according to size on 1% agarose gel electrophoresis, and DNA bands were visualized by UV transillumination after ethidium bromide staining. The same products were also analyzed by electrophoresis on a 5% acrylamide, 2.7 M urea gel (0.1 x 35 x 11 cm, 36 wells) set up in TBE buffer and run at 10 V/cm for 4.5 hours; bands were revealed with silver staining (Sambrook and Russell 2001). A mixture of pUC18 Hae III and lambda phage EcoR I + Hind III digests, which provides a continuous ladder between 2 and 0.08 kb, was used as molecular weight standard.

Restriction fragment analysis

PCR products were screened independently using nine restriction endonucleases (Cfo I, Rsa I, Hae III, Msp I, Dde I, Hinf I, Tru 9I, Taq I and Alu I) each having a 4-bp specific recognition site. Without further purification, 1 μ L of PCR product (about 50 ng) was completely digested for 2-3 hours at 37°C in a final volume of 10 μ L with 1 unit of restriction enzyme (Promega) using supplier's buffer. The reaction was stopped by adding 2.5 μ L of 100 mM EDTA, pH 8.3, containing saturated sucrose and a trace of bromophenol blue. Samples were loaded on an acrylamide-urea gel for an electrophoresis running at 11 V/cm until the bromophenol blue dye reached the end of the gel; bands were revealed with silver staining.

Heteroduplex assays (HDA)

HDA were performed on Dde I restriction fragments according to Sambrook and Russell (2001). Equal amounts of Dde I products from two different sources were mixed, raised and kept at 98°C for 2 min, cooled rapidly to 78°C and slowly brought back to 25°C over a 20 min period. HDA products were analyzed on acrylamide-urea gels as above.

Results

Each of the 11 morphological species (see Table 3.1), with at least two individuals per species being tested, yielded a single PCR product on agarose gel (Fig 3.1A), with sizes distributed between 4 (maybe 5, a small size difference exist between *C. mutata* and *S. quebecense*) discrete levels, at ~ 820, ~ 890, ~ 975 and ~ 1150 bp respectively. A second, often weaker, neighbouring band of a slightly larger size was also seen on polyacrylamide gels (Fig. 3.1B) for about half the species tested, which indicated some sequence/length heterogeneity within the ITS-1 family of repeated units. Six of the 11 morphological species, all from the *Simulium* genus, namely *S. aureum*, *S. venustum*, *S. jenningsi*, *S. decorum*, *S. vittatum* and *S. longistylatum* shared a common major band at ~ 820 bp. At best, the last two could be distinguished from the others through the absence of the minor band at ~ 875 bp, but, given the small number of specimens from each species, it is doubtful whether this trait is characteristic and reproducible for taxonomic purposes. With no intermediate profile being found, two of the specimens morphologically identified as *S. decorum s.l.* yielded a single band around 880-890 bp. This clearly distinguished them from the above *S. decorum*, which exhibited a major/minor profile at 820 and 880-890 bp respectively. For the sake of clarity, this molecular variant, likely another member of the *S. decorum* complex (Adler and McCreadie 1997), will further be designated as *S. decorum**. Including *S. decorum** provisionally considered as species or sibling, the six species outside the 820 bp subgroup clustered on polyacrylamide-urea gels at the same three levels as previously

defined on agarose gels, namely *S. decorum**/*C. dacotensis* at ~ 890 bp, *S. quebecense*/*C. mutata* at ~ 1020 bp and *P. fontanum* / *P. fuscum* at ~ 1230 bp.

From the above results, size differences between PCR products could not alone distinguish all 12 species from each other. Additional data being provided through restriction analysis (Figs. 3.2 and 3.3) fell into three classes. First, some restriction fingerprints added no more information than already available through ITS-1 size (data not shown). For instance, Hae III did not possess any target site within ITS-1 and left it untouched. Cfo I and Msp I each produced a common fragment for all species, plus 1-2 bands with a pattern no more informative than ITS-1 size alone. Next, Taq I slightly improved the classification only for *C. mutata* and *S. quebecense*, since the DNA from the latter was cut, but not for the former.

Finally, Alu I, Rsa I, Tru 91, Hinf I and Dde I fingerprints offered the maximum discriminatory power. They all clearly distinguished *S. decorum** from *C. dacotensis* and *C. mutata* from *S. quebecense* (Figs. 3.2A and 3.2B), which subdivides the 12 species into six sub-groups. Yet, the six *Simulium* species corresponding to the 820 bp level on agarose gels remain unresolved, except for *S. aureum* which, upon digestion with Rsa I, gave an additional band at 108 bp (Fig. 3.2C). This extra band could well arise from the above mentioned spacer length heterogeneity. Indeed, only in the case of *S. aureum*, did the minor size variant of the intact ITS-1 represent close to half the total amount of the material being analysed on polyacrylamide gels. Although some other

species from the *Simulium* genus also possessed such a band, it amounted only to traces, such that no extra fragment could be detected in Rsa I digests.

Within the nine restriction enzymes under investigation, Cfo I, Tru 9I, Msp I, Dde I and Rsa I, in addition to variable length fragments identifying the species signature, also produced, within minor size differences, 1-2 bands which can be qualified as common to all 12 species. The size of the common fragment(s) averaged at 480 (Msp I), 440 (Rsa I), $250 + 80 + 90$ (Dde I), 280 (Cfo I) and $130 + 120$ (Tru 9I) bp. This suggests that a strongly conserved sequence segment within ITS-1, maybe half of the whole spacer length, is shared between 12 species from three different genera. This would leave only 1 to 3 remaining bands as the species signature. On the whole, our data revealed much less available restriction polymorphism than normally hoped for, which likely indicates a low level of molecular diversity in black flies.

As shown by the *C. dacotensis* / *S. decorum** pair, a size identity of ITS-1 could be misleading with respect to classification at the genus level, since, as exemplified by the Alu I fingerprint, *S. decorum** seemed more closely related to the six *Simulium* species having ITS-1 banding at 820 bp. Size variants could also occur if some members of the family undergo local deletions or tandem duplications, followed by the spreading of the mutation throughout the family. This would explain the frequent occurrence of ITS-1 size heterogeneity, as well as the origin of the *S. decorum** variant. In fact, the minor ITS-1 variant from *S. decorum* seen on polyacrylamide gels corresponded in size with the single band observed for *S. decorum**, suggesting that they could be sequence-related. If in this case, restriction digests from *S. decorum* should yield a double array of

bands: stronger ones representing fragments common to both versions of ITS-1 and weaker ones which, either are specific only for the major *S. decorum* variant or, in the case of the minor variant, coincide with those seen in *S. decorum**. Indeed, Dde I fingerprints from *S. decorum* and *S. decorum** shared bands at 248, 93 and 82 bp, but *S. decorum* also showed at 400 bp another one which, although quite strong too, nevertheless appeared weaker than required by molar yield and was not found in *S. decorum** (Fig. 3.2A). In addition, the *S. decorum* profile possessed two weak and closely spaced bands around 460 bp, that is, at the same position as the one found at molar yield level in *S. decorum**. Similar observations were also recorded with other enzymes (data not shown). Generally speaking, other species having a major/minor profile for the entire spacer length also often exhibited minor bands in restriction patterns, but no detailed interpretation was attempted since data could not be compared to an external reference, contrary to the case of the *S. decorum* and *S. decorum** pair.

We analyzed Dde I digests by HDA, in order to differentiate between *S. venustum*, *S. decorum*, *S. jenningsi*, *S. vittatum* and *S. longistylatum*, a task which none of the previous techniques could perform; *S. aureum* was used as the external standard relative to the other species under investigation (Fig. 3.3). Working with restriction fragments rather than with the whole spacer, it was hoped that local sequence diversity would be more apparent, in addition to studying the size range where polyacrylamide gels perform best. Up to approximately 600 bp, the electrophoresis profile corresponded to the simple sum of the two individual patterns, except that the single band normally seen in the above six *Simulium* species around 245 bp was now shadowed by a companion at ~ 260

bp, which likely represented the interspecific hybrid. Surprisingly, while no molar yield bands with sizes over 460 bp appeared in intraspecific controls, all heteroduplex assays showed one or two additional bands located much higher in the gel and having apparent molecular weights comprised between 980 and 1400 bp, double to triple the size of the next lower band on the gel. On the basis of these unusual structures, all the members of the above group of five species, except for *S. venustum/decorum*, are now distinguished. Large size hybrids were also observed with either *P. fontanum* or *P. fuscum* in assays involving the *S. aureum* external standard, but in this case, it did not allow us to distinguish the former from the latter (data not shown). For each of these two unresolved pairs, both members of the pair appear to belong to the same group within the same genus (Xiong and Kocher 1993, Adler et al. 2004). Based on presently available data, we propose the two-step classification scheme illustrated by Fig. 3.4. ITS-1 is first sorted according to size, either 820 bp, or everything else. Next, a Dde I digestion is performed and, according to the previous criterion, either characterized through HDA or analysed without further treatment.

Discussion

This work on black flies from southern Québec was initiated with the ultimate goal of setting up a simple and universal molecular key relying on as few markers as possible. ITS-1 was consequently chosen as a plausible starting PCR target for both theoretical and practical reasons. However, in spite of previous studies suggesting that size differences of ITS-1 could separate closely related black fly species (Brockhouse *et al.* 1993, Tang *et al.* 1996), we observed only a single band on agarose gels for each of the 12 species investigated here, with little size diversity. It was therefore useful to obtain sequence-related information through restriction digests. Contrary to our expectations, groups first defined according to ITS-1 size proved very hard to split into distinct species even by combining information from nine restriction patterns. The frequent sharing between all species of common fragments, from 1 to 3 depending on the restriction enzyme being used, even suggests that most, if not all, of the sequence variability resides within a limited region of ITS-1. With respect to the cluster of five *Simulium* species (excluding *S. aureum*, see results) having an ITS-1 size of 820 bp, we obtained a total of nine typed restriction sites, which represents a cumulative target sequence of 36 bases, about 3.5% of ITS-1 length, being absolutely identical between them. If this figure is typical of the interspecific variability of the whole ITS-1, it would indicate very little sequence divergence within this group. Therefore, it appears fruitless to further increase the pool of restriction enzymes in order to find discriminating restriction sites which are likely to be inexistent. This low variability could well be related to the fact that ITS-1, as part of a multigene family, undergoes correction

processes which limit its rate of fixation. Typing of spacers between single-copy, but highly conserved, genes could well represent a more viable alternative.

At the end, only HDA of Dde I products brought us close to fully differentiating between five species all belonging to the *Simulium* genus. Surprisingly, we then detected an interspecific DNA hybrid having 2-3 times the expected size. If indeed most interspecific sequence differences are clustered within a limited and restriction site-free part of ITS-1, they will show up in only one of the HDA products within which they will create a large unpaired region. Through this region, this forked molecule would then associate with the complementary segments of its HDA counterpart, creating a pseudo-recombinant Holliday-type junction (insert to Fig. 3.3).

In contrast, when performed on a group of 34 tabanid species from Southern Manitoba (Iranpour *et al.* 2004), amplification of the whole ITS yielded in most cases a wide range of amplicon sizes and numbers, and subsequent restriction profiling easily identified each member of the group. Moreover, these authors sometimes even detected some intraspecific variation. Thus, in spite of black flies as a whole being generally considered of more ancient origins than tabanids (Moulton 2000), there appears to be much less ITS-1 sequence divergence within our random sampling of black fly species from Québec. At the same time, it seems unlikely that the fixation rate for mutations hitting in the ITS-1 family of sequences would differ so much between the two groups. However, this apparent discrepancy fits well with the classification of Adler *et al.* (2004), who noticed only a small proportion of species being endemic simultaneously in North and Central/South America. He observed much closer similarities concerning species from

subarctic regions of Eurasia and North America respectively, but a near total lack of identity between the temperate areas of these two regions. Following this, we hypothesize that the actual range of species from Southern Québec originates from a very small group of subarctic founder “generalist” species (Adler *et al.* 2004) which, following the end of the Wisconsin ice age 12,000 years ago, progressively adapted to more temperate conditions and started only recently to diversify into “specialists” having each its own peculiar ecological niche even in the same stream.

The present work represents the first step towards a universal molecular key for the routine identification of black flies. At the same time, one would normally expect that, short of sequencing, ITS-1 alone will lose part of its discriminatory power as other species are progressively added to the actual data. Clearly, as the case with morphological keys, a pool of different criteria will be required for taxonomy purposes. Following this line of thought, we are currently screening other possible targets, each characterized by conserved DNA sequences flanking a variable region.

Acknowledgements

We are grateful to the Université du Québec à Trois-Rivières who provided C. Gaudreau's scholarship and funding during her research. We also thank Christian Back for the identification of *S. jenningsi*.

References

- Adler, P.H., D.C. Currie, and D.M. Wood. 2004. The blackflies (Simuliidae) of North America. ROM publication in sciences. New York, USA.
- Adler, P.H., and J.W. Mc Creadie. 1997. The hidden ecology of black flies: sibling species and ecological scale. American Entomologist, Fall 1997: 153-161.
- Arnheim, N. 1983. Concerted evolution of multigene families in M. Nei, and R.K. Koehn [eds.], Evolution of genes and proteins. Sinauer, Boston Mass. USA.
- Awise, J.C., J. Arnold, R.M. Ball, E. Bermingham, T. Lamb, J.E. Neigel, C.A. Reeb, and N.C. Saunders. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annual Review of Ecology and Systematics 18: 489-522.
- Ballinger-Crabtree, M.E., W.C. IV Black, and B.R. Miller. 1992. Use of genetic polymorphisms detected by the random-amplified polymorphic DNA polymerase chain reaction (RAPD-PCR) for differentiation and identification of *Aedes aegypti* subspecies and population. American Journal of Tropical Medicine and Hygiene 47: 893-901.

Brockhouse, C.L., C.G. Vajime, R. Marin, and R.M. Tanguay, R.M. 1993. Molecular identification of onchocerciasis vector sibling species in black flies (Diptera: Simuliidae). *Biochemical and Biophysical Research Communications* 194: 628-634.

Brower, A.V.Z., and R. DeSalle. 1994. Practical and theoretical considerations for choice of a DNA sequence region in insect molecular systematics, with a short review of published studies using nuclear gene regions. *Annals of the Entomological Society of America* 87: 702-716.

Carew, M.E., V. Pettigrove, and A.A. Hoffmann. 2003. Identifying chironomids (Diptera: Chironomidae) for biological monitoring with PCR-RFLP. *Bulletin of Entomological Research* 93: 483-490.

Caterino, M.S., S. Cho, and F.A.H. Sperling. 2000. The current state of insect molecular systematics : a thriving Tower of Babel. *Annual Review of Entomology* 45: 1-54.

Crosskey, R.W. 1990. The natural history of blackflies. John Wiley & sons Inc., New York.

de Sousa, G.B., G.P. Panzetta de Dutari, and C.N. Gardenal. 1999. Genetic structure of *Aedes albifasciatus* (Diptera: Culicidae) populations in central Argentina determined by random amplified polymorphic DNA-polymerase chain reaction markers. *Journal of Medical Entomology* 36: 400-404.

Dover, G.A. 1986. Molecular drive in multigene families: how biological novelties arise, spread and are assimilated. *Trends in Genetics* 2: 159-165.

Fredeen, F.J.H. 1977. A review of the economic importance of black flies (Diptera: Simuliidae) in Canada. *Quaestiones Entomologicae* 13: 219-229.

Fredeen, F.J.H. 1985. Some economic effects of outbreaks of black flies (*Simulium luggeri* Nicholson and Mickel) in Saskatchewan. *Quaestiones Entomologicae* 21: 175-208.

Iranpour, M., A.M. Schurko, G.R. Klassen, and T.D. Galloway. 2004. DNA fingerprinting of tabanids (Diptera: Tabanidae) and their respective egg masses using PCR- restriction fragment profiling. *The Canadian Entomologist* 136: 605-619.

Krüger, A., A. Gelhaus, and R. Garms. 2000. Molecular identification and phylogeny of East African *Simulium damnosum* S.l. and their relationship with West African species of the complex (Diptera: Simuliidae). *Insect Molecular Biology* 9: 101-108.

Léry, X., B. LaRue, J. Cossette, and G. Charpentier. 2003. Characterization and authentication of insect cell lines using RAPD markers. *Insect Biochemistry and Molecular Biology* 33: 1035-1041.

Linton, Y.-M., A.J. Mordue (Luntz), R.H. Cruickshank, R. Meiswinkel, P.S. Mellor, and J.F. Dallas. 2002. Phylogenetic analysis of the mitochondrial cytochrome oxidase

subunit 1 gene of five species of the *Culicoides imicola* species complex. Medical and veterinary Entomology 16: 139-146.

Manguin, S., R.C. Wilkerson, J.E. Conn, Y. Rubio-Palis, J.A. Danoff-Burg, and D.R. Roberts. 1999. Population structure of the primary malaria vector in South America, *Anopheles darlingi*, using isozyme, random amplified polymorphic DNA, internal transcribed spacer 2, and morphological markers. American Journal of Tropical Medicine and Hygiene 60: 364-376.

Miller, L.J., P.G. Alisopp, G.C. Graham, and D.K. Yeates. 1999. Identification of morphologically similar canegrubs (Coleoptera:Scarabaeidae:Melolonthini) using a molecular diagnostic technique. Australian Journal of Entomology 38: 189-196.

Morales-Hojas, R., R.J. Post, R.A. Cheke, and M.D. Wilson. 2002. Assessment of rDNA IGS as a molecular marker in the *Simulium damnosum* complex. Medical and Veterinary Entomology 16: 395-403.

Moulton, J.K. 2000. Molecular sequence data resolves basal divergences within Simuliidae (Diptera). Systematic Entomology 25: 95-113

Muraji, M., and S. Nakahara. 2002. Discrimination among pest species of *Bactrocera* (Diptera: Tephritidae) based on PCR-RFLP of the mitochondrial DNA. Applied Entomology and Zoology 37: 437-446.

Post, R.J., P.K. Flook, and A.L. Millett. 1993. Methods for preservation of insects for DNA studies. *Biochemical Systematics and Ecology* 21: 85-92.

Pruess, K.P., B.J. Adams, T.J. Parsons, X. Zhu, and T.O. Powers. 2000. Utility of the mitochondrial cytochrome oxidase II gene for resolving relationships among black flies (Diptera: Simuliidae). *Molecular Phylogenetics and Evolution* 16: 286-295.

Ross, K.G., D.D. Shoemaker, M.J. Krieger, C.J. DeHeer, and L. Keller. 1999. Assessing genetic structure with multiple classes of molecular markers: a case study involving the introduced fire ant *Solenopsis invicta*. *Molecular Biology and Evolution* 16: 525-543.

Rothfels, K.H., and D. Featherston. 1981. The population structure of *Simulium vittatum* (Zett.): the IHL-1 and IS-7 sibling species. *Canadian Journal of Zoology* 59: 1857-1883.

Salazar, M., C. Theoduloz, V. Vega, F. Poblete, E. Gonzalez, R. Badilla, and L. Meza-Basso. 2002. PCR-RFLP identification of the Chilean species of *Rhagoletis* (Diptera: Tephritidae) attacking Solanaceae. *Bulletin of Entomological Research* 92: 337-341.

Sambrook, J., and D.W. Russell. 2001. *Molecular cloning: a laboratory manual*, 3rd ed. Cold Spring Harbor Laboratory Press, New York.

Satta, Y., N. Toyonara, C. Okhtaka, Y. Tatsuno, T.K. Watanabe, E.T. Matsuura, S.L. Chigusa, and N. Takahata. 1988. Dubious maternal inheritance of mitochondrial DNA in

Drosophila simulans and the evolution of *Drosophila mauritiana*. Genetic Research 52: 1-6.

Sebastiani, F., R. Meiswinkel, L.M. Gomulski, C.R. Guglielmino, P.S. Mellor, A.R. Malacrida, and G. Gasperi. 2001. Molecular differentiation of the old world *Culicoides imicola* species complex (Diptera: Ceratopogonidae), inferred using random amplified polymorphic DNA markers. Molecular Ecology 10: 1773-1786.

Shouche, Y.S., and M.S. Patole. 2000. Sequence analysis of mitochondrial 16S ribosomal RNA fragment from seven mosquito species. Journal of Biosciences 25: 361-366.

Strachan, T., D.A. Webb, and G.A. Dover. 1985. Transition stages of molecular drive in multi-copy DNA families in *Drosophila*. European Molecular Biology Organization Journal 4: 1701-1708.

Tang, J., L. Toè, C. Back, and T.R. Unnasch. 1996. Intra-specific heterogeneity of the rDNA internal transcribed spacer in the *Simulium damnosum* (Diptera: Simuliidae) complex. Molecular Biology and Evolution 13: 244-252.

Tang, J., J.K. Moulton, K. Pruess, E.W. Cupp, and T.R. Unnasch. 1998. Genetic variation in North American black flies in the subgenus *Psilopelmia* (Simulium: Diptera: Simuliidae). Canadian Journal of Zoology 76: 205-211.

Wood, D.M., B.V. Peterson, D.M. Davies, and H. Gyorkos. 1963. The Black flies (Diptera: Simuliidae) of Ontario. Part II. Larval identification, with description and illustration. *Proceedings of the Entomological Society of Ontario* 93: 99-129.

Xiong, B., and T.D. Kocher. 1993. Phylogeny of sibling species of *Simulium venustum* and *S. verecundum* (Diptera: Simuliidae) based on sequences of the mitochondrial 16S rRNA gene. *Molecular Phylogenetics and Evolution* 2: 293-303.

Zhu, X., K.P. Pruess, and T.O. Powers. 1998. Mitochondrial DNA polymorphism in a black fly, *Simulium vittatum* (Diptera: Simuliidae). *Canadian Journal of Zoology* 76: 440-447.

Table 3.1: Black fly material used in the study.

Species	Collection site	Geographic location
<i>Cnephia dacotensis</i> (Dyar & Shannon)	St-Élie, Qc	46°29' 72°58'
<i>Simulium venustum</i> (Say)	Mauricie area, Qc	NA
<i>Prosimulium fuscum</i> (Symes & Davies)	Mauricie area, Qc	NA
<i>Simulium quebecense</i> (Twinn)	Lac Souris, Qc	46°35' 72°58'
<i>Simulium decorum</i> (Walker)	Lac Souris, Qc	46°35' 72°58'
<i>Simulium vittatum</i> (Zetterstedt)	Lac Souris, Qc	46°35' 72°58'
<i>Cnephia mutata</i> (Malloch)	Lac Souris, Qc	46°35' 72°58'
<i>Prosimulium fontanum</i> (Syme & Davies)	Lac Souris, Qc	46°35' 72°58'
<i>Simulium longistylatum</i> (Shewell)	Chute Dunbar, Qc	46°57' 73°07'
<i>Simulium aureum</i> (Fries)	L'Assomption, Qc	45°50' 73°25'
<i>Simulium jenningsi</i> (Malloch)	Rivière Rideau, Ottawa, ONT	45°27' 75°42'

Fig 3.1: ITS-1 profile on agarose (A) and polyacrylamide (B) gels. Samples as follows: λ EcoRI + Hind III digest (a and n); *S. jenningsi* (b); *S. aureum* (c); *S. venustum* (d); *P. fuscum* (e); *S. decorum* (f); *S. decorum** (g); *S. vittatum* (h); *P. fontanum* (i); *C. mutata* (j); *S. quebecense* (k); *C. dacotensis* (l) and *S. longistylatum* (m).

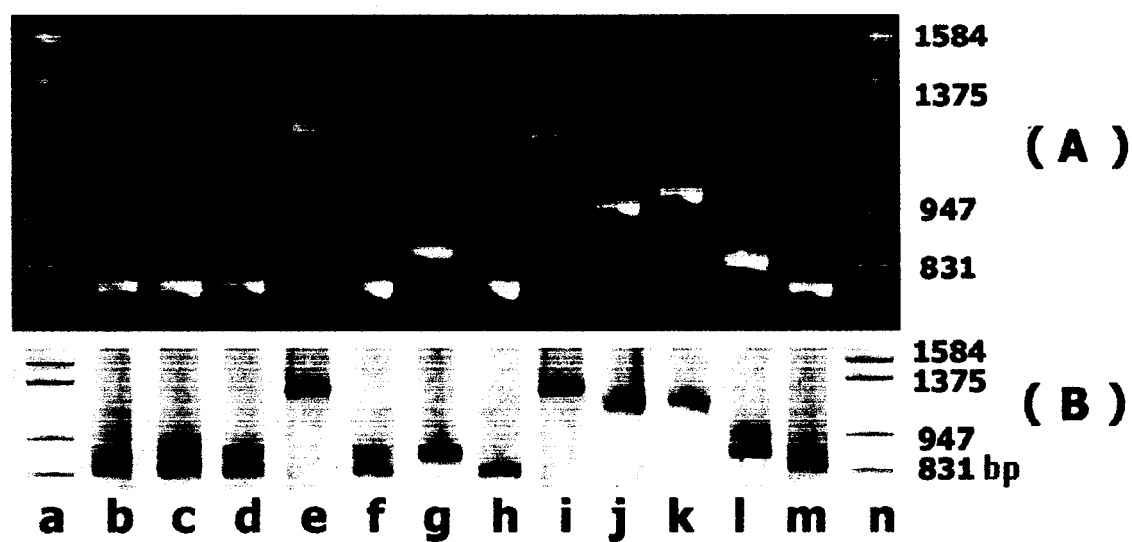


Fig 3.2: Dde I (A), Alu I (B) and Rsa I (C) fingerprints. Samples are: mixture of λ EcoRI + Hind III and pUC18 Hae III digests (a, i and o); *S. longistylatum* (b); *C. dacotensis* (c); *S. quebecense* (d); *C. mutata* (e); *P. fontanum* (f); *S. vittatum* (g); *S. decorum** (h); *S. decorum* (j); *P. fuscum* (k); *S. venustum* (l); *S. aureum* (m) and *S. jenningsi* (n). On the Rsa I profile (C), samples are: *S. venustum* (a); *S. aureum* (b, c and d); *S. jenningsi* (e and f) and mixture of λ EcoRI + Hind III and pUC18 Hae III digests (g).

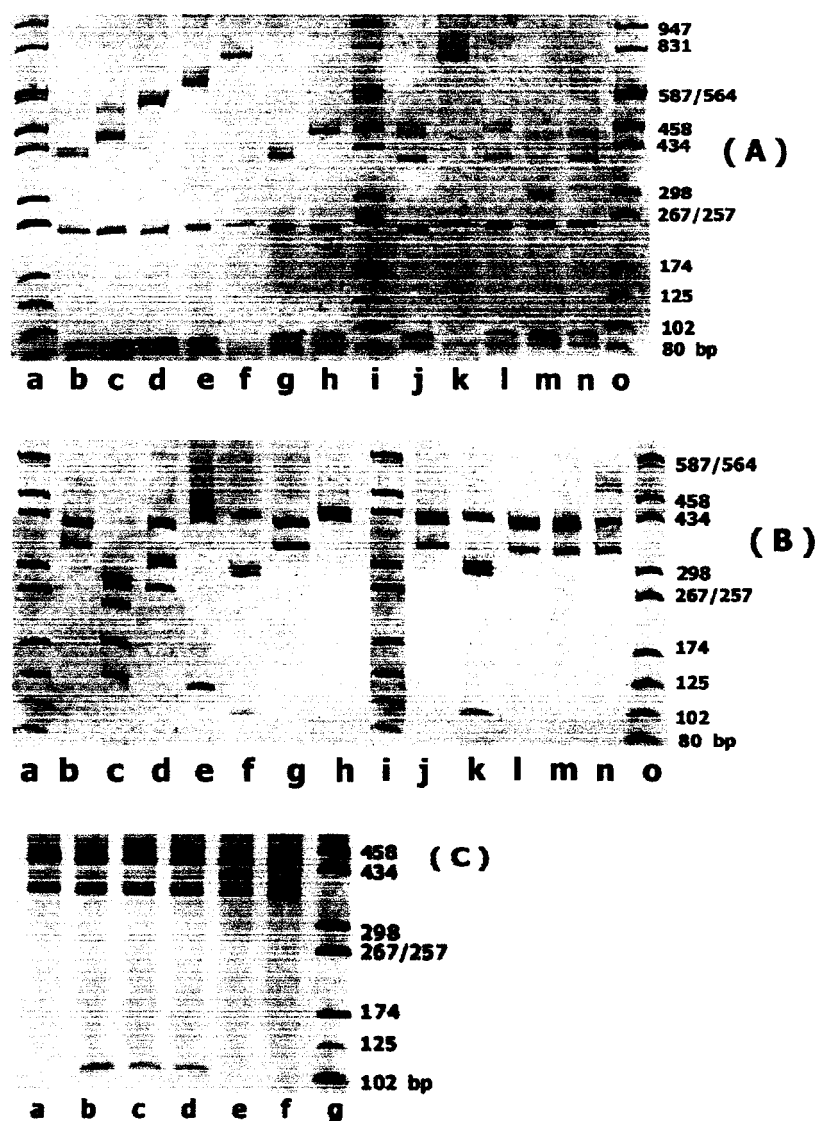


Fig 3.3: HDA profiles from Dde I digest of ITS-1. Samples as follows: mixture of λ EcoRI + Hind III and pUC18 Hae III digests (a, g and n); from b to f, heteroduplexes of *S. jenningsi*, *S. venustum*, *S. decorum*, *S. vittatum* and *S. longistylatum* respectively with *S. aureum*; from h to l, homoduplexes of the same; homoduplex from *S. aureum* (m). Insert: unusually large heteroduplexes and, at right, their tentative structural interpretation.

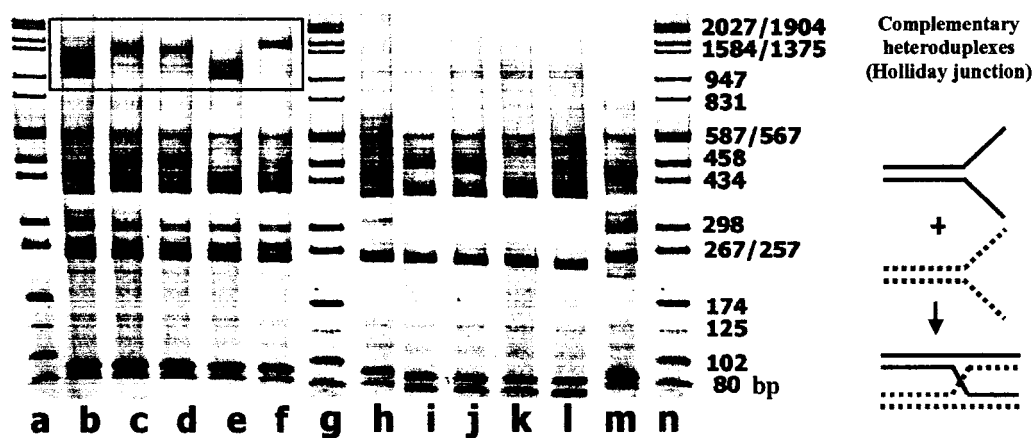
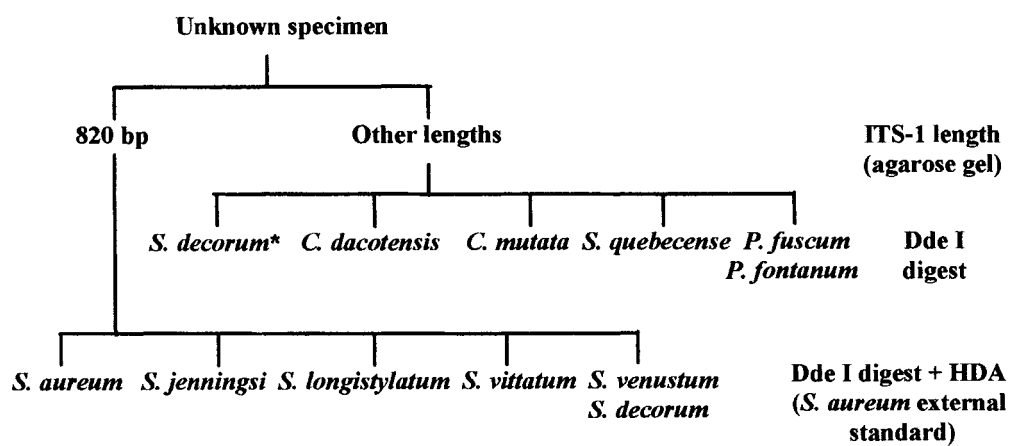


Fig 3.4: Pathway for the classification of 12 black fly species.



ANNEXE I

Nucleic Acids Research submission guidelines

GENERAL REQUIREMENTS

English

Manuscripts must be clearly and concisely written in English. The Editors reserve the right to reject without review those that cannot adequately be assessed because of a poor standard of English. Authors whose first language is not English are encouraged to have their manuscript checked by a native English speaker. Although NAR is unable to recommend an English language editing service for authors, such companies may be identified through search engines. In the event of difficulty in finding a company, authors should contact one of the Senior Editorial Offices (see below).

Abstract

The abstract should be a single paragraph, not exceeding 200 words. URLs and references to figures or schemes should NOT be included. However, note that URLs MUST be included in the abstract of manuscripts submitted to the Database and Web Server issues. A literature reference, especially to work preceding the submission, is permitted. This should have the format Schmidt *et al.* (1995), *Nucleic Acids Res.*, **23**, 4793-4798.

References

These should be cited in the text by sequential number only, in order of appearance, and listed numerically in the References section. Online references should be cited as in example 5, below. Please see examples 6 and 7 for papers that have been published online in more than one version. The initial version of a paper published in this way can be cited by the Digital Object Identifier (doi) but, if available, the reference should also include the citation of the final version. Authors should check all references carefully, and in particular ensure that all references in the Reference section are cited in the text. Note that multiple references or page spans under one number are not allowed. Personal communications, unpublished results, manuscripts submitted or in preparation, statistical packages, computer programs and web sites should be cited in the text only; NOT included in the References section. Accession numbers may be cited either within the text or in the form of a reference

Citations should conform to the following examples. Journal names should be abbreviated in the style of *Chemical Abstracts*. Where the list of authors is extensive it is acceptable to list the first 10 authors followed by *et al.* **NOTE THAT FULL TITLES OF JOURNAL ARTICLES MUST BE PROVIDED.**

1. Schmitt,E., Panvert,M., Blanquet,S. and Mechulam,Y. (1995) Transition state stabilisation by the 'high' motif of class I aminoacyl-tRNA synthetases: the case of *Escherichia coli* methionyl-tRNA synthetase. *Nucleic Acids Res.*, **23**, 4793-4798.
2. Huynh,T.V., Young,R.A. and Davies,R.W. (1988) Constructing and screening cDNA libraries in lambda gt10 and lambda gt11. In Glover,D.M. (ed.), *DNA Cloning - A Practical Approach*. IRL Press, Oxford, Vol. I, pp. 49-78.
3. Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

4. Burnett,R.C. (1993) EMBL accession no. X52486.
5. Capaldi,S., Getts,R.C. and Jayasena,S.D. (2000) Signal amplification through nucleotide extension and excision on a dendritic DNA platform. *Nucleic Acids Res.*, **28**, e21.
6. Qiao,D., Chen,W., Stratagoules,E. and Martinez,J. (March 10, 2000) Bile acid-induced activation of activator protein-1 requires both extracellular signal-regulated kinase and protein kinase C signaling. *J. Biol. Chem.*, 10.1074/jbc.M908890199
7. Qiao,D., Chen,W., Stratagoules,E. and Martinez,J. (2000) Bile acid-induced activation of activator protein-1 requires both extracellular signal-regulated kinase and protein kinase C signaling. *J. Biol. Chem.*, **275**, 15090-15098. First published on May 19, 2000, 10.1074/jbc.M908890199
8. Bernhagen,J., Elkin,B., Geiger,G., Tovar,G. and Vitzthum,F. (1999) Patent DE-198198889.2-44; PCT/WO/EP/99/03047.

Nomenclature conventions

Restriction enzymes, DNA methyltransferases and homing endonucleases should be named or referred to using the conventions described in Roberts,R.J. *et al.* (2003) *Nucleic Acids Res.*, **31**, 1805-1812. Note that restriction enzyme names should NOT be italicized.

Figures

Upon **acceptance** you will be required to submit high-resolution images (figures and schemes) direct to the production office at Oxford University Press (contact details below). Therefore, it is important that you prepare high-resolution source files in the first instance. Low-resolution images for the review process can then be prepared from these source files.

For production, electronic images must be at a minimum resolution of 600 d.p.i. for line drawings (black and white) and 300 d.p.i. for colour or greyscale. Colour figures must be supplied in CMYK not RGB colours. Please ensure that the prepared electronic image files print at a legible size (with lettering of at least 2 mm) and are of a high quality for publication.

For useful information on preparing figures visit <http://dx.sheridan.com> where you can also test whether your figures are suitable for production by using the proflight tool at <http://dx.sheridan.com/onl/>.

For your **initial submission**, please prepare a second version of your figures at low resolution (150 d.p.i.); this will be used in the reviewing process and will ensure that editors and reviewers will avoid long file downloading times. It is preferable that you insert these low-resolution figures and tables into the word processing document, after the text. If you are unable to do this the figure files may be uploaded separately, in .jpg, .gif, .tif or .eps format. If possible, please combine your figures into a single file. When inserting figures into your text documents, please make sure they are readable; many figures contain tiny characters (such as numbers on a chart or graph), and if these characters are not easily readable in your text document, they will most likely be illegible in the .pdf prepared by the system. Each figure must be clearly numbered.

- **Size and layout**

The area of a page is 230 mm (height); 84 mm (single) or 178 mm (double) (column width). Figures should not exceed these dimensions and ideally should fit either a single or double column. Lettering should be of a consistent size within each figure.

- **Colour figures**

The use of colour in figures may improve clarity in some cases. Although there are no formal colour charges, the use of colour is subject to editorial discretion. If an Executive Editor or the Production office recommend a colour figure to be printed in black and white and the author disagrees, the author will be offered the option to pay for the colour, otherwise it will be printed in black and white.

- **Figure legends**

Symbols must NOT be included in figure legends; ALL symbols used in figures should be described in the legend using words (e.g. black triangle, open like to submit a cover image for consideration, please click [here](#) for details).

ANNEXE II

Journal of Medical entomology submission guidelines

General Submission Instructions

All manuscripts submitted to the Entomological Society of America should be written with clarity and readability in mind. Manuscripts are subject to editing to ensure conformity to editorial standards and journal style. Consult the Council of Biology Editors' style manual, *Scientific Style and Format Manual for Authors, Editors, and Publishers*, 6th ed., along with recent issues of the journal for style and format. Follow the requirements shown below when preparing manuscripts and refer to specific ESA editorial standards as noted in the [ESA Style Guide](#).

- Submit manuscript as an MS Word, WordPerfect, or RTF file with a page size of Letter, 8.5 x 11".
- Type all as double-spaced, with 1-inch margins, and do not right justify text.
- Use the font Times (New) Roman with a size of 12 point.
- Left-justify the title, author line, affiliation lines, subheadings, text, and References Cited.
- Insert tabs, not spaces, for paragraph indents.
- Use italicization only to indicate scientific names (including viruses), symbols or variables, and words that are defined.
- Use quotation marks for quoted material only.
- Use American English spelling throughout and follow *Merriam-Webster's New Collegiate Dictionary*, 10th ed., for guidance on spelling.
- Number pages consecutively, beginning with the title page.
- Use continuous line numbers on all text pages.
- Begin each of the following on a separate page and arrange in the following order: title page, abstract and key words (three to six words), text, acknowledgments, references cited, footnotes, tables, figure legends, and figures.
- Type all captions on a separate page and put each figure and table on a separate page.

Refer to the [Figures](#) section of the ESA Style Guide for information on how to submit figures.

Refer to the [ESA Style Guide](#) for additional information.